

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 13:53:18 ; Search time 2407.08 Seconds
(without alignments)
234.084 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKPWMPWRK 13

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09444281/runat_07052004_171138_2568/app_query.fasta_1.398
-DB=GenEmbl -QWMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09444281 @CNG 1.1 3608 @runat_07052004_171138_2568 -NCFU=6 -ICFU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	91	100.0	114	6	BD247517	BD247517 Method fo
2	76	83.5	108	6	BD247523	BD247523 Method fo
3	76	83.5	114	6	BD247529	BD247529 Method fo
4	76	83.5	151	6	BD247520	BD247520 Method fo
5	73	80.2	39	6	BD128612	BD128612 Method fo
6	73	80.2	68	6	AX357080	AX357080 Sequence
7	73	80.2	69	6	AR404925	AR404925 Sequence
8	73	80.2	211	6	AR226388	AR226388 Sequence
9	73	80.2	211	6	AR282754	AR282754 Sequence
10	73	80.2	550	4	BTINDICD	X67340 Bos taurus
11	73	80.2	6446	6	AX173324	AX173324 Sequence
12	73	80.2	6446	6	AX098418	AX098418 Sequence
13	69	75.8	156785	2	AC146517	AC146517 Homo sapi
14	69	75.8	157518	9	AC093525	AC093525 Homo sapi
15	69	75.8	176552	9	AC106820	AC106820 Homo sapi
16	69	75.8	205268	2	AC146518	AC146518 Homo sapi
17	68	74.7	141040	8	AP004382	AP004382 Oryza sat
18	67	73.6	195764	2	AC113900	AC113900 Rattus no
19	67	73.6	228163	2	AC120671	AC120671 Rattus no
20	67	73.6	241568	2	AC129646	AC129646 Rattus no
21	67	73.6	255121	2	AC095460	AC095460 Rattus no
22	66	72.5	207	6	BD273664	BD273664 Peptides
23	66	72.5	2651	10	MMU08210	U08210 Mus musculu
24	66	72.5	3643	10	BC051649	BC051649 Mus muscu
25	66	72.5	65209	2	AC101355	AC101355 Mus muscu
26	66	72.5	107257	10	AF289665	AF289665 Mus muscu
27	66	72.5	146408	8	AC130610	AC130610 Oryza sat
28	66	72.5	175494	2	AC127242	AC127242 Mus muscu
29	66	72.5	179914	10	AC124511	AC124511 Mus muscu
30	66	72.5	201395	10	AC091250	AC091250 Mus muscu
31	65	71.4	7505	1	CAC300832	AJ300832 Pelittia a
32	65	71.4	27780	2	AC020365	AC020365 Drosophil
33	65	71.4	110000	2	BX255276.07	Continuation (8 of
34	65	71.4	149991	10	AL672300	AL672300 Mouse DNA
35	65	71.4	160817	3	AC008316	AC008316 Drosophil
36	65	71.4	177028	3	AC008315	AC008315 Drosophil
37	65	71.4	233552	2	AC131774	AC131774 Mus muscu
38	65	71.4	238323	3	AE003684	AE003684 Drosophil
39	65	71.4	269759	2	AC093467	AC093467 Mus muscu
40	65	71.4	301332	1	AE017012	AE017012 Bacillus
41	64	70.3	39	6	BD273655	BD273655 Peptides
42	64	70.3	53	6	BD273657	BD273657 Peptides
43	64	70.3	54	6	BD273659	BD273659 Peptides
44	64	70.3	2718	10	BC029697	BC029697 Mus muscu
45	64	70.3	19479	10	NMA2IXCOA	Z22923 M.musculus

ALIGNMENTS

RESULT 1

```

BD247517      114 bp      DNA      linear      PAT 17-JUL-2003
LOCUS          Method for effectively producing antibacterial cationic peptides in
DEFINITION     host cells.
ACCESSION      BD247517
VERSION        BD247517.1 GI:33057287
KEYWORDS       JP 2002530114-A/11.
SOURCE         synthetic construct
ORGANISM       artificial construct
               artificial sequences.
REFERENCE      1 (bases 1 to 114)
AUTHORS       Burian,J. and Bartfeld,D.
TITLE         Method for effectively producing antibacterial cationic peptides in
JOURNAL       host cells
PATENT        JP 2002530114-A 11 17-SEP-2002;
COMMENT       MICROLOGIX BIOTECH INC
PN            JP 2002530114-A/11
PD            17-SEP-2002
PF            19-NOV-1999 JP 2000584088
PR            20-NOV-1998 US 60/109218
PI            JAN BURIAN,DANIEL BARTFELD
PC            C12N15/09,C07K1/12,C07K1/18,C12N1/15,C12N1/19,C12N1/21,C12P21/
PC            02//
CC            Synthesized oligonucleotide used as a template for PCR FH
KEY           Location/Qualifiers
FT            source 1..114
FT            /db_xref="taxon:32630"
FEATURES       source
               Location/Qualifiers
               1..114
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
ORIGIN
Alignment Scores:
Pred. No.:      0.00429      Length:      114
Score:          91.00       Matches:    13
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00%     Indels:    0
DB:             6           Gaps:     0

US-09-444-281-35 (1-13) x BD247517 (1-114)
QY            1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
Db            ::::
Db            41 ATCCGAAAAAATGGCGGTGGCGGTGGCGGTGCGTAA 79

RESULT 2
BD247523
LOCUS          108 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION     Method for effectively producing antibacterial cationic peptides in
ACCESSION      BD247523
VERSION        BD247523.1 GI:33057293
KEYWORDS       JP 2002530114-A/17.
SOURCE         synthetic construct
ORGANISM       artificial construct
               artificial sequences.
REFERENCE      1 (bases 1 to 108)
AUTHORS       Burian,J. and Bartfeld,D.
TITLE         Method for effectively producing antibacterial cationic peptides in
JOURNAL       host cells
PATENT        JP 2002530114-A 17 17-SEP-2002;
COMMENT       MICROLOGIX BIOTECH INC
PN            JP 2002530114-A/17
PD            17-SEP-2002
PF            19-NOV-1999 JP 2000584088
PR            20-NOV-1998 US 60/109218
PI            JAN BURIAN,DANIEL BARTFELD

```

```

PC            C12N15/09,C07K1/12,C07K1/18,C12N1/15,C12N1/19,C12N1/21,C12P21/
PC            02//
PC            (C12P21/02,C12R1/19),C12N15/00
CC            Synthesized oligonucleotide us as a template for PCR FH Key
               Location/Qualifiers
               1..108
               /organism='Artificial Sequence'.
FEATURES       source
               Location/Qualifiers
               1..108
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
ORIGIN
Alignment Scores:
Pred. No.:      0.247      Length:      108
Score:          76.00       Matches:    9
Percent Similarity: 92.31%   Conservative: 3
Best Local Similarity: 69.23% Mismatches: 1
Query Match:    83.52%     Indels:    0
DB:             6           Gaps:     0

US-09-444-281-35 (1-13) x BD247523 (1-108)
QY            1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
Db            ::::
Db            38 ATGATCTCGGTGGCGGTGGCGGTGGCGGTGCGCAA 76

RESULT 3
BD247529
LOCUS          114 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION     Method for effectively producing antibacterial cationic peptides in
ACCESSION      BD247529
VERSION        BD247529.1 GI:33057299
KEYWORDS       JP 2002530114-A/23.
SOURCE         synthetic construct
ORGANISM       artificial construct
               artificial sequences.
REFERENCE      1 (bases 1 to 114)
AUTHORS       Burian,J. and Bartfeld,D.
TITLE         Method for effectively producing antibacterial cationic peptides in
JOURNAL       host cells
PATENT        JP 2002530114-A 23 17-SEP-2002;
COMMENT       MICROLOGIX BIOTECH INC
PN            JP 2002530114-A/23
PD            17-SEP-2002
PF            19-NOV-1999 JP 2000584088
PR            20-NOV-1998 US 60/109218
PI            JAN BURIAN,DANIEL BARTFELD
PC            C12N15/09,C07K1/12,C07K1/18,C12N1/15,C12N1/19,C12N1/21,C12P21/
PC            02//
CC            (C12P21/02,C12R1/19),C12N15/00
CC            Synthesized oligonucleotide used as a template for PCR FH
KEY           Location/Qualifiers
FT            source 1..114
FT            /organism='Artificial Sequence'.
FEATURES       source
               Location/Qualifiers
               1..114
               /organism="synthetic construct"
               /mol_type="genomic DNA"
               /db_xref="taxon:32630"
ORIGIN
Alignment Scores:
Pred. No.:      0.262      Length:      114
Score:          76.00       Matches:    9
Percent Similarity: 92.31%   Conservative: 3
Best Local Similarity: 69.23% Mismatches: 1
Query Match:    83.52%     Indels:    0
DB:             6           Gaps:     0

```



```
RESULT 7
AR404925 AR404925 69 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6630197.
ACCESSION AR404925
VERSION AR404925.1 GI:40153719
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 69)
AUTHORS Wood, T. K., Jayaraman, A. and Earthman, J. C.
TITLE Inhibition of sulfate-reducing-bacteria-mediated degradation using
bacteria which secrete antimicrobials
JOURNAL Patent: US 6630197-A 1 07-OCT-2003;
FEATURES Location/Qualifiers
source 1..69
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 0.352 Length: 69
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AR404925 (1-69)

QY 4 LysTrpProTrpTrpProTrpArgArg 12
Db 28 AAATGGCTTGGTGGCGCTGGCGCGC 54

RESULT 8
AR226388 AR226388 211 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 11 from patent US 644645.
ACCESSION AR226388
VERSION AR226388.1 GI:27264888
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 211)
AUTHORS Selsted, M. E. and Osapay, K.
TITLE Crosslink-stabilized indolicidin analogs
JOURNAL Patent: US 644645-A 11 03-SEP-2002;
FEATURES Location/Qualifiers
source 1..211
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.13 Length: 211
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AR226388 (1-211)

QY 4 LysTrpProTrpTrpProTrpArgArg 12
Db 38 AAATGGCGCTGGTGGCGCTGGCGCGT 64

RESULT 9
AR282754 AR282754 211 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 38 from patent US 6524585.
```

```
ACCESSION AR282754 GI:29719494
VERSION AR282754.1
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 211)
AUTHORS Selsted, M. E.
TITLE Indolicidin analogs and methods of using same
JOURNAL Patent: US 6524585-A 38 25-FEB-2003;
FEATURES Location/Qualifiers
source 1..211
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.13 Length: 211
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AR282754 (1-211)

QY 4 LysTrpProTrpTrpProTrpArgArg 12
Db 38 AAATGGCGCTGGTGGCGCTGGCGCGT 64

RESULT 10
BTINDLCD BTINDLCD 550 bp mRNA linear MAM 15-NOV-2001
LOCUS Bos taurus mRNA for cathelicidin (CATHL4 gene).
DEFINITION X67340
ACCESSION X67340.1 GI:462
KEYWORDS cathelicidin; CATHL4 gene; indolicidin antimicrobial peptide.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1
AUTHORS Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M.
TITLE cDNA cloning of the neutrophil bactericidal peptide indolicidin
JOURNAL Biochem. Biophys. Res. Commun. 187 (1), 467-472 (1992)
MEDLINE 92392368
PUBMED 1520337
REFERENCE 2 (bases 1 to 550)
AUTHORS Del Sal, G.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1992) G. Del Sal, Univ. of Trieste, Dip. di
Biochimica, Biofisica e Chimica, delle Macromolecole, Via A.
Valerio, 38, 34127 Trieste, ITALY
FEATURES Location/Qualifiers
source 1..550
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/cell_line="bone marrow cells"
1..550
/gene="CATHL4"
13..447
/gene="CATHL4"
/function="indolicidin antimicrobial peptide"
/codon_start=1
/product="cathelicidin"
/protein_id="CAA47755.1"
/db_xref="GI:463"
/db_xref="GOA:P33046"
/db_xref="SWISS-PROT:P33046"
/translation="MQQTRASLIGRWLWLLGLVVPSSAQALSREAVLRAVDQ
INELSSLEANLYKLELDFPPKUNEDIGTRKPVSTVKTVCPTTIQQFAEQCDFKNG
```

RVKQCVGTWILDPSNDFDINCNELQSVLPKWPWRRG"
 13..99
 /gene="CATHL4"
 403..444
 /gene="CATHL4"
 /product="cathelicidin"

ORIGIN

Alignment Scores:
 Pred. No.: 3.08 Length: 550
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x BTINDLCD (1-550)

Qy 4 LysTirProTirPirProTirPirArg 12
 Db 415 AATGGCCATGGTGGCTGGCGGAGA 441

RESULT 11

AR173324
 LOCUS AR173324 6446 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 5 from patent US 6303779.
 ACCESSION AR173324
 VERSION AR173324.1 GI:17912815
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 6446)
 AUTHORS Garger,S.J., Holtz,R.Barry., McCulloch,M.J. and Turpen,T.H.
 TITLE Processes for isolating and purifying viruses and sugars from plant sources

JOURNAL Patent: US 6303779-A 5 16-OCT-2001;

FEATURES
 source

1..6446
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 40.3 Length: 6446
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AR173324 (1-6446)

Qy 4 LysTirProTirPirProTirPirArg 12
 Db 6213 AATGGCCATGGTGGCCATGGCGGAGA 6239

RESULT 12

AX098418
 LOCUS AX098418 6446 bp RNA linear PAT 03-APR-2001
 DEFINITION Sequence 5 from Patent WO0119969.
 ACCESSION AX098418
 VERSION AX098418.1 GI:13537710

KEYWORDS
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiids; Solanales; Solanaceae; Nicotiana.

REFERENCE

1
 AUTHORS Garger,S.J., Holtz,B.R., McCulloch,M.J. and Turpen,T.H.
 TITLE A process for isolating and purifying viruses, soluble proteins and peptides from plant sources

JOURNAL Patent: WO 0119969-A 5 22-MAR-2001;
 Large Scale Biology Corporation (US)
 FEATURES
 source
 Location/Qualifiers
 1..6446
 /organism="Nicotiana benthamiana"
 /mol_type="unassigned RNA"
 /db_xref="taxon:4100"

ORIGIN

Alignment Scores:
 Pred. No.: 40.3 Length: 6446
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AX098418 (1-6446)

Qy 4 LysTirProTirPirProTirPirArg 12
 Db 6213 AATGGCCATGGTGGCCATGGCGGAGA 6239

RESULT 13

AC146517/c
 LOCUS AC146517 156785 bp DNA linear HTG 24-AUG-2003
 DEFINITION Homo sapiens chromosome 16 clone RP11-67B18, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.

ACCESSION AC146517
 VERSION AC146517.1 GI:34147313
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 156785)
 DOB Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished

2 (bases 1 to 156785)
 DOB Joint Genome Institute.

REFERENCE Direct Submission

TITLE Submitted (24-AUG-2003) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT

-----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 432135
 Center clone name: RPCI-11_67B18

Summary Statistics

Consensus quality: 147563 bases at least Q40
 Consensus quality: 149669 bases at least Q30
 Consensus quality: 150963 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 155385; sum-of-contigs estimation
 Quality coverage: 9.57 in Q20 bases; agarose-fp estimation
 Quality coverage: 10.77 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1118: contig of 1118 bp in length
 * 1119 1218: gap of unknown length
 * 1219 2329: contig of 1111 bp in length
 * 2330 2429: gap of unknown length

* 2430 3567: contig of 1138 bp in length
 * 3568 3667: gap of unknown length
 * 3668 4716: contig of 1049 bp in length
 * 4717 4816: gap of unknown length
 * 4817 7812: contig of 2996 bp in length
 * 7813 7912: gap of unknown length
 * 7913 14712: contig of 6800 bp in length
 * 14713 14812: gap of unknown length
 * 14813 19661: contig of 4849 bp in length
 * 19662 19761: gap of unknown length
 * 19762 24662: contig of 4901 bp in length
 * 24663 24762: gap of unknown length
 * 24763 34140: contig of 9378 bp in length
 * 34141 34240: gap of unknown length
 * 34241 43353: contig of 9113 bp in length
 * 43354 43453: gap of unknown length
 * 43454 56333: contig of 12880 bp in length
 * 56334 56433: gap of unknown length
 * 56434 74223: contig of 17790 bp in length
 * 74224 74323: gap of unknown length
 * 74324 90148: contig of 15825 bp in length
 * 90149 90248: gap of unknown length
 * 90249 114156: contig of 23908 bp in length
 * 114157 114256: gap of unknown length
 * 114257 156785: contig of 42529 bp in length.

FEATURES Location/Qualifiers
 source 1..156785
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-67B18"
 /clone_lib="RPC1 human BAC library 11"

ORIGIN

Alignment Scores:
 Pred. No.: 3.38e+03 Length: 156785
 Score: 69.00 Matches: 10
 Percent Similarity: 91.67% Conservative: 1
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 75.82% Indels: 0
 DB: 2 Gaps: 0

US-09-444-281-35 (1-13) x AC146517 (1-156785)
 QY 2 LeuLysLysTrpProTTrpProTTrpArgArgLys 13
 |||:|||||
 DB 8135 TTACAAAATCCCTCGTGGCCCTGGAGGAGGAG 8100

RESULT 14
 AC093525
 LOCUS Homo sapiens chromosome 16 clone RP11-20123, complete sequence.
 DEFINITION
 AC093525
 AC093525.3 GI:26023955
 VERSION
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 157518)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 157518)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 157518)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.

TITLE Direct Submission
 JOURNAL Submitted (04-DEC-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT On Dec 4, 2002 this sequence version replaced gi:16041346.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence;
 Estimated Total Number of Errors is 0.2.

FEATURES

source 1..157518
 Location/Qualifiers
 1..157518
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-20123"

ORIGIN

Alignment Scores:
 Pred. No.: 3.4e+03 Length: 157518
 Score: 69.00 Matches: 10
 Percent Similarity: 91.67% Conservative: 1
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 75.82% Indels: 0
 DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AC093525 (1-157518)

QY 2 LeuLysLysTrpProTTrpProTTrpArgArgLys 13
 |||:|||||
 DB 106227 TTACAAAATCCCTCGTGGCCCTGGAGGAGGAG 106262

RESULT 15

AC106820
 LOCUS Homo sapiens chromosome 16 clone RP11-715J22, complete sequence.
 DEFINITION
 AC106820
 AC106820.4 GI:38524627
 VERSION
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE 1 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (12-JAN-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 4 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUL-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 5 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 6 (bases 1 to 176552)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.

COMMENT

On Nov 26, 2003 this sequence version replaced gi:21672112.
 Draft Sequence Produced by DOE Joint Genome Institute

www.igi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.4.

FEATURES

Location/Qualifiers
1. .176552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-715J22"

ORIGIN

Alignment Scores:
Pred. No.: 3.83e+03 Length: 176552
Score: 69.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 75.82% Indels: 0
DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AC106820 (1-176552)

QY 2 LeuLysLysTrpProTrrProTrrArgArgLys 13
|||::|||
Db 16954 TTCAAAATCCCTGGTGGCCCTGGAGGAGGAG 16989

Search completed: May 11, 2004, 16:01:29
Job time : 2429.08 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 12:39:07 ; Search time 256.88 Seconds

(without alignments)
214.990 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKXWPPWPRK 13

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame*_p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09444281/runat_07052004_171138_2561/app.query.fasta_1.398
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORSE=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281_@C@N_1_1_586@runat_07052004_171138_2561 -NCFU=6 -ICFU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	91	100.0	88	2	AAV60908
2	91	100.0	114	3	AAA27291
3	76	83.5	108	3	AAA27296
4	76	83.5	114	3	AAA27298
5	76	83.5	151	3	AAA27294
6	73	80.2	39	2	AAV83788
7	73	80.2	47	3	AAZ29389
8	73	80.2	47	3	AAZ29390

c	9	73	80.2	47	6	ABL60414
	10	73	80.2	47	6	ABL60415
	11	73	80.2	53	3	AAZ29364
	12	73	80.2	53	6	ABL60445
	13	73	80.2	68	6	AAI72481
	14	73	80.2	69	3	AAZ40246
	15	73	80.2	211	3	AAZ49764
	16	73	80.2	211	3	AAZ45123
	17	73	80.2	550	6	AAZ45350
	18	73	80.2	6446	2	AAZ20646
	19	73	80.2	6446	2	AAZ20646
	20	66	72.5	207	3	AAA28519
	21	65	71.4	1848	4	ABL21537
	22	65	71.4	4433	4	ABL21536
	23	64	70.3	39	3	AAA28510
	24	64	70.3	39	9	ADC73335
	25	64	70.3	53	9	ADC73337
	26	64	70.3	54	9	ADC73339
	27	63	69.2	1040	6	ABL58680
	28	63	69.2	1040	6	ABA96630
	29	63	69.2	1437	4	AAZ54129
	30	63	69.2	1908	4	AAZ56079
	31	63	69.2	1908	7	ACA51647
	32	63	69.2	6957	9	ADB52723
	33	62	68.1	48	3	AAA27290
	34	62	68.1	49	3	AAA27287
	35	62	68.1	52	3	AAA28514
	36	62	68.1	52	3	AAA28512
	37	62	68.1	701	7	ABX94441
	38	62	68.1	1029	4	AAF60750
	39	61.5	67.6	1539	7	ADA70241
	40	61	67.0	2213	4	ABL22542
	41	61	67.0	125439	6	ABQ88177
	42	60	65.9	4031	2	AAZ84485
	43	60	65.9	14739	9	ADC86792
	44	60	65.9	34570	8	ADA02891
	45	60	65.9	34570	9	ADB72629

ALIGNMENTS

RESULT 1
AAV60908
ID AAV60908 standard; DNA; 88 BP.
XX
AC AAV60908;
XX
DT 11-JAN-1999 (first entry)
XX
DE DNA fragment encoding MB111.
XX
KW MB128; cationic peptide; plasmid pXL1; small cryptic plasmid;
XX replication; RepA; vector; RAMP; human; MB111; ss.
XX
OS Synthetic.
XX Homo sapiens.
XX
PN W09841636-A2.
XX
PD 24-SEP-1998.
XX
PF 16-MAR-1998; 98WO-CA000214.
XX
PR 14-MAR-1997; 97US-0040722P.
XX
PA (BURI/) BURIAN J.
XX (KAYW/) KAY W W.
XX
PI Burian J, Kay WW;
XX
DR WPI; 1998-531571/45.
XX
PT Increasing plasmid copy number in a cell with the repA gene product - and

PT an small cryptic plasmid ori sequence, useful for high level expression
 PT of e.g. cytokines, antigens or therapeutic proteins.
 XX Example 16; Page 57; 82pp; English.

XX This oligonucleotide was used as a template in a PCR reaction (see also
 CC AAV60909-10) to generate a DNA fragment encoding the cationic peptide
 CC MB111 (see AAW71690). The PCR product was cloned into the universal
 CC vector pR2h-B1, which contains the R21 replication leader of RepA (see
 CC AAW71686) and 2 tandem copies of the prepro region (hpro) of human
 CC defensin. The vector provides expression of R21-hpro-MB111 fusion in host
 CC cells. The invention provides controlled replication plasmid vectors
 CC (RAMP vectors) comprising a replication origin of a small cryptic plasmid
 CC such as pK11 (see AAV58292) and a gene encoding RepA (see AAW71686). The
 CC vectors can reach very high levels of plasmid replication, but are not
 CC lethal to the host cell, and can be used to direct the high level
 CC expression of e.g. cytokines, antigens and therapeutic proteins
 XX Sequence 88 BP; 20 A; 18 C; 25 G; 25 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.00318 Length: 88
 Score: 91.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-444-281-35 (1-13) x AAV60908 (1-88)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 Db 25 ATCTCGAAAAAATGGCGTGTGGCGTGTGGCGTGTAA 63

RESULT 2

AAA27291
 ID AAA27291 standard; DNA; 114 BP.

AC AAA27291;

DT 20-SEP-2000 (first entry)

DE Oligonucleotide used for synthesis of MBI-11 fragment.

XX Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.

XX Synthetic.

XX WO200031279-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products.

XX Example 4; Page 37; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the

CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is an
 CC oligonucleotide that was used to synthesise a MBI-11 fragment. MBI-11 is
 CC a cationic peptide derived from modifications of indolicidin
 XX Sequence 114 BP; 25 A; 26 C; 30 G; 33 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.00421 Length: 114
 Score: 91.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAA27291 (1-114)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 Db 41 ATCTCGAAAAAATGGCGTGTGGCGTGTGGCGTGTAA 79

RESULT 3

AAA27296
 ID AAA27296 standard; DNA; 108 BP.

XX AAA27296;

XX 20-SEP-2000 (first entry)

XX Oligonucleotide used for synthesis of MBI 2X11B7 poly cassette.

XX Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.

XX Synthetic.

XX WO200031279-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products.

XX Example 5; Page 39; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is an
 CC oligonucleotide that was used to synthesise a MBI-11B7 fragment. This
 CC fragment was used in the expression cassette. MBI-11B7 is a cationic
 CC peptide derived from modifications of indolicidin

XX Sequence 108 BP; 18 A; 33 C; 31 G; 26 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.251 Length: 108.
 Score: 76.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3
 Best Local Similarity: 69.23% Mismatches: 1
 Query Match: 83.52% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAA27296 (1-108)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 AAA27298
 Db 38 ATGATTCTGCGTGGCCGTGGTGGCCGTGGCGTGCACAA 76

RESULT 4
 AAA27298
 ID AAA27298 standard; DNA; 114 BP.
 AC AAA27298;
 XX
 XX 20-SEP-2000 (first entry)
 DT
 DE Oligonucleotide used for synthesis of MBI 11B7 first cassette.

XX Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.
 XX
 OS Synthetic.
 XX
 PN WO200031279-A2.
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA001107.
 XX
 PR 20-NOV-1998; 98US-0109218P.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

PI Burian J, Bartfeld D;
 DR WPI; 2000-400086/34.
 XX
 XX Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products.
 PS Example 5; Page 40; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is an
 CC oligonucleotide that was used to synthesise a MBI-11B7 fragment. This
 CC fragment was used in the expression cassette. MBI-11B7 is a cationic
 CC peptide derived from modifications of indolicidin
 XX
 SQ Sequence 114 BP; 20 A; 34 C; 32 G; 28 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.266 Length: 114
 Score: 76.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3
 Best Local Similarity: 69.23% Mismatches: 1
 Query Match: 83.52% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAA27298 (1-114)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 AAA27298
 Db 38 ATGATTCTGCGTGGCCGTGGTGGCCGTGGCGTGCACAA 76

RESULT 4
 AAA27298
 ID AAA27298 standard; DNA; 114 BP.
 AC AAA27298;
 XX
 XX 20-SEP-2000 (first entry)
 DT

Oligonucleotide used for synthesis of MBI 11B7 last cassette.
 DE Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.
 XX
 OS Synthetic.
 XX
 PN WO200031279-A2.
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA001107.
 XX
 PR 20-NOV-1998; 98US-0109218P.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

PI Burian J, Bartfeld D;
 DR WPI; 2000-400086/34.
 XX
 XX Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products.
 PS Example 5; Page 38; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is an
 CC oligonucleotide that was used to synthesise a MBI-11B7 fragment. This
 CC fragment was used in the expression cassette. MBI-11B7 is a cationic
 CC peptide derived from modifications of indolicidin
 XX
 SQ Sequence 151 BP; 22 A; 44 C; 49 G; 36 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.36 Length: 151
 Score: 76.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3
 Best Local Similarity: 69.23% Mismatches: 1
 Query Match: 83.52% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAA27294 (1-151)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 AAA27294
 Db 38 ATGATTCTGCGTGGCCGTGGTGGCCGTGGCGTGCACAA 76

RESULT 6
 AAA27294
 ID AAA27294 standard; DNA; 151 BP.
 AC AAA27294;
 XX
 XX 20-SEP-2000 (first entry)
 DT

Oligonucleotide used for synthesis of MBI 2X11B7 last cassette.
 DE Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.
 XX
 OS Synthetic.
 XX
 PN WO200031279-A2.
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA001107.
 XX
 PR 20-NOV-1998; 98US-0109218P.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

PI Burian J, Bartfeld D;
 DR WPI; 2000-400086/34.
 XX
 XX Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products.
 PS Example 5; Page 38; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a multi-
 CC domain fusion protein expression cassette comprising a promoter and a
 CC nucleic acid molecule expressed as an insoluble protein. The inclusion of
 CC anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the fusion
 CC protein is controlled. This cassette allows high yield, stable production
 CC of the cationic peptide. Cationic peptides such as bovine indolicidin may
 CC be used as antimicrobial agents. The present sequence is an
 CC oligonucleotide that was used to synthesise a MBI-11B7 fragment. This
 CC fragment was used in the expression cassette. MBI-11B7 is a cationic
 CC peptide derived from modifications of indolicidin
 XX
 SQ Sequence 151 BP; 22 A; 44 C; 49 G; 36 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.36 Length: 151
 Score: 76.00 Matches: 9
 Percent Similarity: 92.31% Conservative: 3
 Best Local Similarity: 69.23% Mismatches: 1
 Query Match: 83.52% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAA27294 (1-151)

QY 1 IleLeuLysLysTrpProTrpTrpProTrpArgArgLys 13
 AAA27294
 Db 38 ATGATTCTGCGTGGCCGTGGTGGCCGTGGCGTGCACAA 76

RESULT 6
 AAA27294
 ID AAA27294 standard; DNA; 151 BP.
 AC AAA27294;
 XX
 XX 20-SEP-2000 (first entry)
 DT

Oligonucleotide used for synthesis of MBI 2X11B7 last cassette.
 DE Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;
 KW indolicidin; bovine; ss.
 XX
 OS Synthetic.
 XX
 PN WO200031279-A2.
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA001107.
 XX
 PR 20-NOV-1998; 98US-0109218P.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

KW purF gene; fusion peptide; mass production; pharmaceutical industry;
 KW food industry; ss.
 XX Synthetic.

OS WO9964611-A1.

PN 16-DEC-1999.

PD 08-JUN-1999; 99WO-KR000282.

PF 09-JUN-1998; 98KR-00022117.

PR 14-MAY-1999; 99KR-00017920.

XX (SAMY-) SAMYANG GENEX CORP.

XX Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;

XX WPI; 2000-097542/08.

XX New DNA constructs useful for mass production of antimicrobial peptides

XX in microorganism hosts.

XX Example 1; Page 13; 67pp; English.

XX The present sequence is a chemically synthesised PCR primer which was

XX used to synthesise a gene encoding antimicrobial peptide Indolicidin. The

XX antimicrobial peptide gene is used in a DNA construct that comprises

XX entire, partial or a derivative of purF gene (glutamine

XX pyrophosphoribosyl pyrophosphate amidotransferase gene). The DNA

XX construct allows mass production of the antimicrobial peptide in

XX microbial hosts without killing the host cells. The antimicrobial

XX peptides are useful commercially in the pharmaceutical and food

XX industries

XX SQ Sequence 47 BP; 12 A; 18 C; 10 G; 7 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.235 Length: 47

Score: 73.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 80.22% Indels: 0

DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AA229390 (1-47)

Qy 4 LysTrpProTrpTrpProTrpArgArg 12

Db 35 AAATGGCCGTGGTGGCGTGGCGTGTGT 9

RESULT 9

ABL60414

ID ABL60414 standard; DNA; 47 BP.

AC ABL60414;

XX 28-MAR-2003 (first entry)

XX DNA fragment of the invention #15.

XX Gene expression; peptide antibiotic; purF gene; ds.

XX Unidentified.

XX KR2001098973-A.

XX 08-NOV-2001.

XX 08-JUN-2001; 2001KR-00031889.

XX 08-JUN-2001; 2001KR-00031889.

(SAMY-) SAMYANG GENEX CORP.

XX Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;

XX WPI; 2002-301977/34.

XX Gene expression system useful for mass-production of peptide antibiotics

XX and vectors derived from microorganisms.

XX Disclosure; Page 7; 56pp; Korean.

XX The invention relates to a genetic component which mass-produces peptide

XX antibiotics effectively from microorganisms. The genetic component

XX consists of a first gene sequence which codes for the whole or partial

XX purF gene or its derivative, and a second gene sequence which codes

XX peptide antibiotics. The mass-production method of peptide antibiotics

XX comprises the steps of; constructing an expression vector including the

XX genetic component, transforming a bacterial host cell with the vector,

XX culturing the transformed cell to express the genetic component, and

XX recovering the peptide antibiotics. The expression vector is selected

XX from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a

XX high copy number of origin, strong transcription promoter and structural

XX gene. The sequences given in records ABL60400-ABL60464 represent DNA

XX sequences of the invention

XX SQ Sequence 47 BP; 6 A; 11 C; 19 G; 11 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.235 Length: 47

Score: 73.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 80.22% Indels: 0

DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x ABL60414 (1-47)

Qy 4 LysTrpProTrpTrpProTrpArgArg 12

Db 17 AAATGGCCGTGGTGGCGTGGCGTGTGT 43

RESULT 10

ABL60415/c

ID ABL60415 standard; DNA; 47 BP.

XX ABL60415;

XX 28-MAR-2003 (first entry)

XX DNA fragment of the invention #16.

XX Gene expression; peptide antibiotic; purF gene; ds.

XX Unidentified.

XX KR2001098973-A.

XX 08-NOV-2001.

XX 08-JUN-2001; 2001KR-00031889.

XX 08-JUN-2001; 2001KR-00031889.

XX (SAMY-) SAMYANG GENEX CORP.

XX Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;

XX WPI; 2002-301977/34.

XX Gene expression system useful for mass-production of peptide antibiotics

XX and vectors derived from microorganisms.

XX Disclosure; Page 7; 56pp; Korean.

XX The invention relates to a genetic component which mass-produces peptide

XX antibiotics effectively from microorganisms. The genetic component

XX consists of a first gene sequence which codes for the whole or partial

XX purF gene or its derivative, and a second gene sequence which codes

XX peptide antibiotics. The mass-production method of peptide antibiotics

XX comprises the steps of; constructing an expression vector including the

XX genetic component, transforming a bacterial host cell with the vector,

XX culturing the transformed cell to express the genetic component, and

XX recovering the peptide antibiotics. The expression vector is selected

XX from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a

XX high copy number of origin, strong transcription promoter and structural

XX gene. The sequences given in records ABL60400-ABL60464 represent DNA

XX sequences of the invention

XX SQ Sequence 47 BP; 6 A; 11 C; 19 G; 11 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.235 Length: 47

Score: 73.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 80.22% Indels: 0

DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x ABL60414 (1-47)

Qy 4 LysTrpProTrpTrpProTrpArgArg 12

Db 17 AAATGGCCGTGGTGGCGTGGCGTGTGT 43

RESULT 10

ABL60415/c

ID ABL60415 standard; DNA; 47 BP.

XX ABL60415;

XX 28-MAR-2003 (first entry)

XX DNA fragment of the invention #16.

XX Gene expression; peptide antibiotic; purF gene; ds.

XX Unidentified.

XX KR2001098973-A.

XX 08-NOV-2001.

XX 08-JUN-2001; 2001KR-00031889.

XX 08-JUN-2001; 2001KR-00031889.

XX (SAMY-) SAMYANG GENEX CORP.

XX Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;

XX WPI; 2002-301977/34.

XX Gene expression system useful for mass-production of peptide antibiotics

XX and vectors derived from microorganisms.

XX Disclosure; Page 7; 56pp; Korean.

XX The invention relates to a genetic component which mass-produces peptide
 CC antibiotics effectively from microorganisms. The genetic component
 CC consists of a first gene sequence which codes for the whole or partial
 CC purF gene or its derivative, and a second gene sequence which codes
 CC peptide antibiotics. The mass-production method of peptide antibiotics
 CC comprises the steps of: constructing an expression vector including the
 CC genetic component, transforming a bacterial host cell with the vector,
 CC culturing the transformed cell to express the genetic component, and
 CC recovering the peptide antibiotics. The expression vector is selected
 CC from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a
 CC high copy number of origin, strong transcription promoter and structural
 CC gene. The sequences given in records ABL60400-ABL60464 represent DNA
 CC sequences of the invention

XX SQ Sequence 47 BP; 12 A; 18 C; 10 G; 7 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.235 Length: 47
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x ABL60415 (1-47)

QY 4 LysTrpProTrpTrpProTrpArgArg 12

DB 35 AAATGGCGGTGGTGGCGGTGGCGTGT 9

RESULT 11

AAZ29364

ID AAZ29364 standard; DNA; 53 BP.

XX AC AAZ29364;

XX DT 29-FEB-2000 (first entry)

XX DE Antimicrobial peptide, Indolicidin encoding DNA.

XX purF gene; glutamine pyrophosphoribosyl pyrophosphate amidotransferase;
 KW purF derivative; fusion partner; antimicrobial peptide; indolicidin;
 KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct; cow;
 KW neutralise; toxicity; pharmaceutical industry; food industry; ds.

XX OS Bos taurus.

XX FH Key Location/Qualifiers

XX FT CDS 5..46

XX FT /*tag= a

XX FT /product= "Indolicidin peptide"

XX FT /note= "Antimicrobial peptide used in DNA construct"

XX PN WO9964611-A1.

XX PD 16-DEC-1999.

XX PF 08-JUN-1999; 99WO-KR000282.

XX PR 09-JUN-1998; 98KR-00022117.

XX PR 14-MAY-1999; 99KR-00017920.

XX PA (SAMY-) SAMYANG GENEX CORP.

XX PI Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;

XX WPI; 2000-097542/08.

XX P-PSDB; AAY44324.

XX New DNA constructs useful for mass production of antimicrobial peptides
 PT in microorganism hosts.

XX

PS Claim 1; Fig 1; 67pp; English.

XX The present DNA sequence encodes an antimicrobial peptide, Indolicidin
 CC derived from cow, Bos taurus. It is used along with a derivative of purF
 CC gene sequence that functions as a fusion partner. A DNA construct that
 CC comprises this antimicrobial peptide encoding sequence and the entire,
 CC partial or derivative of purF gene, is used for mass production of the
 CC antimicrobial peptide in microorganisms without killing the host cells.
 CC Use of the purF gene derivative sequence, neutralises the toxicity of the
 CC antimicrobial peptides against the host microorganism. The antimicrobial
 CC peptides are useful commercially in the pharmaceutical and food
 CC industries

XX SQ Sequence 53 BP; 8 A; 12 C; 20 G; 13 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.267 Length: 53
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x AAZ29364 (1-53)

QY 4 LysTrpProTrpTrpProTrpArgArg 12

DB 17 AAATGGCGGTGGTGGCGGTGGCGTGT 43

RESULT 12

ABL60445

ID ABL60445 standard; DNA; 53 BP.

XX AC ABL60445;

XX DT 28-MAR-2003 (first entry)

XX DE DNA fragment of the invention #44.

XX KW Gene expression; peptide antibiotic; purF gene; ds.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT CDS 5..60

XX FT /*tag= a

XX FT /partial

XX FT /note= "no start codon present"

XX PN KR2001098973-A.

XX PD 08-NOV-2001.

XX PF 08-JUN-2001; 2001KR-00031889.

XX PR 08-JUN-2001; 2001KR-00031889.

XX PA (SAMY-) SAMYANG GENEX CORP.

XX PI Hong SS, Kang MH, Kim JH, Lee HS, Lee JW, Park SH;

XX WPI; 2002-301977/34.

XX P-PSDB; ABB81940.

XX Gene expression system useful for mass-production of peptide antibiotics
 PT and vectors derived from microorganisms.

XX PS Disclosure; Page 17; 56pp; Korean.

XX The invention relates to a genetic component which mass-produces peptide
 CC antibiotics effectively from microorganisms. The genetic component
 CC consists of a first gene sequence which codes for the whole or partial
 CC purF gene or its derivative, and a second gene sequence which codes

CC peptide antibiotics. The mass-production method of peptide antibiotics
 CC comprises the steps of: constructing an expression vector including the
 CC genetic component, transforming a bacterial host cell with the vector,
 CC culturing the transformed cell to express the genetic component, and
 CC recovering the peptide antibiotics. The expression vector is selected
 CC from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a
 CC high copy number of origin, strong transcription promoter and structural
 CC gene. The sequences given in records ABL60400-ABL60464 represent DNA
 CC sequences of the invention

XX Sequence 53 BP; 8 A; 12 C; 20 G; 13 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.267 Length: 53
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x ABL60445 (1-53)

QY 4 LysTrpProTrpTrpProTrpArgArg 12
 DB 17 AAATGGCGGTGGTGGCGTGGCGTGGT 43

RESULT 13

AAI72481
 ID AAI72481 standard; cDNA; 68 BP.

XX AC AAI72481;

DT 16-MAY-2002 (first entry)

XX cDNA derived from C-terminus of ubiquitin.

XX Recursive ensemble mutagenesis; REM; cell viability; optical signal;
 KW high-throughput screening; antimicrobial compound; antibiotic; ss.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..60

FT /*tag= a

FT /partial

FT /transl_except= pos:28..30, aa:Trp

FT /note= "No start codon given"

XX WO200206517-A2.

XX PD 24-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US023004.

XX PR 19-JUL-2000; 2000US-0219179P.

XX PA (KAIR-) KAIROS SCI INC.

XX BYlina EJ, Coleman WJ, Youvan DC;

XX WPI; 2002-179801/23.

XX P-PSDB; MABA7907.

XX Screening compounds affecting cell viability e.g. for identifying
 PT antimicrobial compounds, comprises determining if induced transformed
 PT cell colonies have a desired signal when contacted with a viability
 PT indicator.

XX Disclosure; Fig 6; 56pp; English.

XX This sequence is derived from the C-terminus of ubiquitin and represents
 CC the cloning region. SacII-BglII cassettes can be used for inserting
 CC antimicrobial peptide sequences into this region. This sequence may be

CC used to form the recursive ensemble mutagenesis (REM) cassette of the
 CC invention. The cassette may be used in the method of the invention for
 CC determining whether a compound affects cell viability by: (a) exposing
 CC colonies of cells (CC) on a support surface to inducing conditions, where
 CC the cells have been transformed with an expression library encoding that
 CC candidate compounds; (b) contacting CC with a viability indicator that
 CC produces an optical signal indicative of cell viability; and (c)
 CC determining if a colony has a desired optical signal. The method is
 CC useful for determining whether a compound affects cell viability. It is
 CC useful for high-throughput screening to identify antimicrobial compounds
 CC and in drug discovery. The antimicrobial compounds are useful in the
 CC pharmaceutical industry, and provide an additional new class of
 CC antibiotic compounds to fight infectious diseases. The method is useful
 CC for assaying the authentic peptide sequences contained in an expression
 CC library for antimicrobial activity, for distinguishing dead cells
 CC (expressing active sequences) from living cells (expressing inactive or
 CC less active sequences), to identify novel antimicrobial peptide
 CC sequences, including highly potent molecules, resulting in a large number
 CC of new antimicrobial lead compounds that are active against a broad range
 CC of bacteria or other microorganisms, and for screening all types of
 CC antibiotic compounds, including libraries of low molecular weight
 CC molecules produced by metabolic engineering and artificial synthesized
 CC libraries in solid-phase arrays

XX Sequence 68 BP; 11 A; 18 C; 19 G; 20 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.349 Length: 68
 Score: 73.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.22% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-35 (1-13) x AAI72481 (1-68)

QY 4 LysTrpProTrpTrpProTrpArgArg 12

DB 31 AAATGGCGTGGTGGCGTGGCGTGGC 57

RESULT 14

AAZ40246

ID AAZ40246 standard; DNA; 69 BP.

XX AC AAZ40246;

DT 23-FEB-2000 (first entry)

XX DE Oligonucleotide for cloning indolicidin peptide coding sequence.

XX Indolicidin; bactenecin; sulphate-reducing bacteria; growth inhibitor;
 KW corrosion; degradation; metal; concrete; cement; dental implant; biofilm;
 KW ss.

XX Synthetic.

OS Bacillus sp.

XX PN WO9556553-A1.

XX PD 11-NOV-1999.

XX PF 03-MAY-1999; 99WO-US009675.

XX PR 06-MAY-1998; 98US-00074037.

XX PR 31-MAR-1999; 99US-00282277.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Wood TK, Jayaraman A, Earthman JC;

XX WPI; 2000-052882/04.

XX Inhibiting growth of sulfate-reducing bacteria using other bacteria,

particularly for protection of metals and concrete.

Example 4; Fig 1; 84pp; English.

This sequence represents an oligonucleotide for cloning the non-amidated indolicidin peptide coding sequence. The invention relates to a method for inhibiting growth of sulphate-reducing bacteria (A) on a material (B) sensitive to corrosion or degradation, by applying to (B) a bacterium (C) that secretes a compound (I) able to inhibit growth of (A). The method is used to protect metal, concrete or cement against corrosion and degradation, but (B) can also be used to protect dental implants. (B) is present in an open or closed system (e.g. water cooling tower, liquid storage container, fuel tank, sewer or drainage system etc.) or part of a bridge or other structure. The method is more effective and less expensive than known methods for inhibiting (A), and reduces the amount of toxic chemicals released. Conventional biofilms of aerobic organisms tend to encourage growth of (A), and addition of (C) to the biofilm prevents this. A single application of (C) lasts for a long time, and (I) are produced exactly where they are required and inhibit (A) without significant impact on other organisms (this effect includes reducing resistance of (A) to conventional biocides, which may then be used in reduced amounts). If local damage to the biofilm occurs, the underlying material is still protected by diffusion of (I) from neighbouring areas

Sequence 69 BP; 14 A; 18 C; 20 G; 17 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.355	Length:	69
Score:	73.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.22%	Indels:	0
DB:	3	Gaps:	0

US-09-444-281-35 (1-13) x AAZ40246 (1-69)

QY 4 LysTrpProTrrPrProTrrArgArg 12

Db 28 AAATGGCCTTGGTGGCCTTGGCGCGC 54

RESULT 15

AAZ49764

ID AAZ49764 standard; DNA; 211 BP.

XX AC AAZ49764;

XX 18-APR-2000 (first entry)

XX DE Poly- (Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)³ DNA.

XX Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13; stability; bovine neutrophil; antimicrobial; antibacterial; fungicide; KX protozoacide; virucide; anti-HIV; human immunodeficiency virus-1; HIV-1; KX gram positive bacteria; gram negative; Staphylococcus aureus; KX Escherichia coli; Salmonella typhimurium; Yeast; fungi; protozoa; KX Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba; KX hexapeptide spacer; ds.

XX OS Synthetic.

XX BOS SP.

XX Key Location/Qualifiers

FT primer_bind 1..21

FT /tag= b

FT CDS 8..199

FT /tag= a

FT /product= "Poly- (Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)³"

FT /note= "encodes three copies of Indol 1-13, each

FT separated by Met-Ala-Arg-Ile-Ala-Met spacer sequence"

FT 68..71

FT /tag= d

FT /note= "corresponds to overlap in oligonucleotides used

FT for ligation"

FT

FT misc_feature 148..151

FT /tag= e

FT /note= "Corresponds to overlap in oligonucleotides used

FT for ligation"

FT primer_bind complement(191..211)

FT /tag= c

FT WO9965510-A1.

FN 23-DEC-1999.

PD 20-MAY-1999; 99WO-US011165.

PF 18-JUN-1998; 98US-00099631.

XX (REGC) UNIV CALIFORNIA.

XX Selsted ME, Osapay K;

PI WPI; 2000-147133/13.

DR P-PSDB; AAY44668.

DR Crosslinked indolicidin analogs with antimicrobial activity against

XX bacteria, yeast, fungi, protozoa and viruses.

XX Example 1C; Fig 1; 53pp; English.

XX The patent discloses crosslinked analogs of indolicidin (Indol 1-13) which is a naturally occurring peptide isolated from bovine neutrophils and has antimicrobial activity. The crosslinked indolicidin (X-indolicidin) analogs are stable and have antimicrobial activity against gram positive and negative bacteria (e.g. Staphylococcus aureus, Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g. Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia species and Acanthamoeba species), and viruses (e.g. HIV-1). They can be used for reducing or inhibiting the growth or survival of microorganisms in an environment e.g. a food or food product, a solution, an inanimate object comprising a surface, or a mammal. The present sequence is a DNA encoding a protein comprising three copies of Indol 1-13 each separated by a hexapeptide spacer sequence. The sequence was used to produce a recombinant construct for the expression of Indol-homoserine (Hse) analog. The ability of Indol-Hse analog to maintain antimicrobial activity provides a means to produce X-indolicidin analog precursors in sufficient quantities

XX SQ Sequence 211 BP; 36 A; 50 C; 74 G; 51 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.18	Length:	211
Score:	73.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.22%	Indels:	0
DB:	3	Gaps:	0

US-09-444-281-35 (1-13) x AAZ49764 (1-211)

QY 4 LysTrpProTrrPrProTrrArgArg 12

Db 38 AAATGGCCTTGGTGGCCTTGGCGCGC 64

Search completed: May 11, 2004, 14:43:51

Job time : 260.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:02:47 ; Search time 49.92 Seconds
(without alignments)
144.518 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKQWPPWPERK 13

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09444281/runat_07052004_171139_2589/app.query.fasta_1.398
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281@cgn2_1_1_76 @runat_07052004_171139_2589 -NCFU=6 -ICFU=3
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	88	3	US-09-042-071-49
2	73	80.2	39	3	Sequence 49, Appl
3	73	80.2	69	4	Sequence 29, Appl
4	73	80.2	211	4	Sequence 1, Appl
5	73	80.2	211	4	Sequence 11, Appl
6	73	80.2	6446	3	Sequence 38, Appl
7	73	80.2	6446	3	Sequence 5, Appl
8	73	80.2	6446	4	Sequence 5, Appl
9	63	69.2	1278	4	Sequence 9269, Ap
10	63	69.2	1362	4	Sequence 9298, Ap
11	63	69.2	1521	4	Sequence 9281, Ap
12	60	65.9	4031	1	Sequence 1, Appl

c	13	58	63.7	292	4	US-09-313-294A-5936	Sequence 5936, Ap
	14	58	63.7	1509	2	US-08-481-337A-1	Sequence 1, Appl
	15	58	63.7	1509	2	US-08-696-268B-1	Sequence 1, Appl
	16	58	63.7	1509	4	US-08-448-371A-1	Sequence 1, Appl
	17	58	63.7	1509	5	PCT-US95-05467-1	Sequence 1, Appl
	18	58	63.7	1596	5	PCT-US94-11328A-3	Sequence 3, Appl
	19	58	63.7	1733	4	US-09-023-655-1402	Sequence 1402, Ap
	20	58	63.7	1970	4	US-09-023-655-1394	Sequence 1394, Ap
	21	58	63.7	1984	3	US-09-382-256-1	Sequence 1, Appl
	22	58	63.7	1984	3	US-09-395-115-1	Sequence 1, Appl
	23	58	63.7	1984	4	US-08-436-265-1	Sequence 1, Appl
	24	58	63.7	1984	4	US-08-679-187-1	Sequence 1, Appl
c	25	58	63.7	111282	4	US-09-754-250-3	Sequence 3, Appl
	26	57.5	63.2	1894	3	US-08-749-816-1	Sequence 1, Appl
	27	57.5	63.2	1894	4	US-09-144-914-1	Sequence 1, Appl
	28	57.5	63.2	7032	2	US-08-149-097D-24	Sequence 24, Appl
	29	57.5	63.2	7032	3	US-08-949-386-24	Sequence 24, Appl
	30	57.5	63.2	7032	3	US-08-450-562-24	Sequence 24, Appl
	31	57.5	63.2	7032	4	US-08-984-709A-24	Sequence 24, Appl
	32	57.5	63.2	7032	4	US-08-450-272-24	Sequence 24, Appl
	33	57.5	63.2	7032	4	US-08-450-273-24	Sequence 24, Appl
	34	57.5	63.2	7089	3	US-08-949-386-25	Sequence 25, Appl
	35	57.5	63.2	7089	3	US-08-450-562-25	Sequence 25, Appl
	36	57.5	63.2	7089	4	US-08-984-709A-25	Sequence 25, Appl
	37	57.5	63.2	7089	4	US-08-450-272-25	Sequence 25, Appl
	38	57.5	63.2	7089	4	US-08-450-273-25	Sequence 25, Appl
	39	57	62.6	3271	4	US-09-548-797B-1	Sequence 1, Appl
	40	57	62.6	3390	4	US-09-548-797B-2	Sequence 2, Appl
	41	57	62.6	3431	4	US-09-632-098-1	Sequence 1, Appl
	42	57	62.6	3468	4	US-09-632-098-3	Sequence 3, Appl
	43	57	62.6	3582	4	US-09-548-797B-3	Sequence 3, Appl
	44	57	62.6	17000	4	US-09-548-797B-7	Sequence 7, Appl
	45	56	61.5	654	4	US-09-280-116-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1

US-09-042-071-49
; Sequence 49, Application US/09042071
; Patent No. 6294372
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; TITLE OF INVENTION: REPLICATION GENES AND GENE PRODUCTS FROM SMALL CRYPTIC PLASMIDS AND METHODS FOR CONSTRUCTING CONTROLLED-REPLICATION PLASMID VECTORS
; TITLE OF INVENTION: SMALL CRYPTIC PLASMIDS AND METHODS FOR CONSTRUCTING CONTROLLED-REPLICATION PLASMID VECTORS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,071
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McWaters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 660081.407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:


```
; LENGTH: 88 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-042-071-49
Alignment Scores:
Pred. No.: 0.00168 Length: 88
Score: 91.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x US-09-042-071-49 (1-88)
QY 1 IieLeuLysLysTrpProTrrpProTrrpArgArgLys 13
Db 25 ATCTGAAATAATGGCGTGGTGGCGTGGCGTGGTAA 63
RESULT 2
US-09-230-180-29
; Sequence 29, Application US/09230180
; Patent No. 6183992
; GENERAL INFORMATION:
; APPLICANT: Kim, Sun-Chang
; APPLICANT: Lee, Jae Hyun
; APPLICANT: Kang, Min Hyung
; APPLICANT: Kim, Jeong Hyun
; APPLICANT: Hong, Seung-Suh
; APPLICANT: Lee, Hyun-Soo
; APPLICANT: Samwang Genex Corporation
; APPLICANT: Korea Advanced Institute of Science and Technology
; TITLE OF INVENTION: METHOD FOR MASS PRODUCTION OF
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDE
; FILE REFERENCE: 6181/0F135
; CURRENT APPLICATION NUMBER: US/09/230.180
; PRIOR FILING DATE: 1999-03-10
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: PCT/KR98/00132
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: KR 21312/1997
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence deduced from Indolicidin peptide
; OTHER INFORMATION: sequence based on codon usage of E. coli
US-09-230-180-29
Alignment Scores:
Pred. No.: 0.0847 Length: 39
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x US-09-230-180-29 (1-39)
QY 4 LysTrpProTrrpProTrrpArgArg 12
Db 13 AAATGGCGTGGTGGCGTGGCGTGGT 39
RESULT 3
US-09-282-277-1
; Sequence 1, Application US/09282277
; Patent No. 6630197
```

```
; GENERAL INFORMATION:
; APPLICANT: Wood, Thomas K.
; APPLICANT: Jayaraman, Arul
; APPLICANT: Earthman, James C.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Inhibition of Sulfate-Reducing-Bacteria-Mediated
; TITLE OF INVENTION: Degradation Using Bacteria Which Secrete Antimicrobials
; FILE REFERENCE: 02307E-085910US
; CURRENT APPLICATION NUMBER: US/09/282.277
; CURRENT FILING DATE: 1999-03-31
; EARLIER APPLICATION NUMBER: US 09/074,037
; EARLIER FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
; OTHER INFORMATION: used for cloning indolicidin
US-09-282-277-1
Alignment Scores:
Pred. No.: 0.153 Length: 69
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x US-09-282-277-1 (1-69)
QY 4 LysTrpProTrrpProTrrpArgArg 12
Db 28 AAATGGCGTGGTGGCGTGGCGTGGCGGC 54
RESULT 4
US-09-099-631A-11
; Sequence 11, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Osapay, Klara
; TITLE OF INVENTION: Crosslink-Stabilized Indolicidin Analogs
; FILE REFERENCE: P-UC 3050
; CURRENT APPLICATION NUMBER: US/09/099,631A
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)...(196)
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-099-631A-11
Alignment Scores:
Pred. No.: 0.491 Length: 211
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x US-09-099-631A-11 (1-211)
QY 4 LysTrpProTrrpProTrrpArgArg 12
```

Db 38 AATGGCCGTGGTGGCGTGGCGTGTGT 64

RESULT 5

US-09-416-481A-38
; Sequence 38, Application US/09416481A
; Patent No. 6524585
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3794
; CURRENT APPLICATION NUMBER: US/09/416,481A
; CURRENT FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 09/076,227
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)..(196)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
US-09-416-481A-38

Alignment Scores: 0.491 Length: 211
Pred. No.: 73.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 80.22% Gaps: 0
DB:

US-09-444-281-35 (1-13) x US-09-416-481A-38 (1-211)

QY 4 LysTTPProTTPProTTPProTTPProTTP 12

Db 38 AATGGCCGTGGTGGCGTGGCGTGTGT 64

RESULT 6

US-09-259-741-5
; Sequence 5, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: McCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751

; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8100
; TELEFAX: 650-463-8400
; TELEX:

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
US-09-259-741-5

Alignment Scores: 17.2 Length: 6446
Pred. No.: 73.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 80.22% Gaps: 0
DB:

US-09-444-281-35 (1-13) x US-09-259-741-5 (1-6446)

QY 4 LysTTPProTTPProTTPProTTPProTTP 12

Db 6213 AAGUGGCCUGUGGUGGCCAUGGCCCGA 6239

RESULT 7

US-09-037-751-5
; Sequence 5, Application US/09037751
; Patent No. 6037456
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: McCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,751
; FILING DATE: 10-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 6446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
US-09-037-751-5

Alignment Scores:
Pred. No.: 17.2 Length: 6446
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 3 Gaps: 0

US-09-444-281-35 (1-13) x US-09-037-751-5 (1-6446)

QY 4 LysTpProTtpTtpProTtpArg 12
DB 6213 AAGUGCCUUGGUGGCGGCGCA 6239

RESULT 8

US-09-466-422-5
Sequence 5, Application US/09466422
Patent No. 6303779
GENERAL INFORMATION:

APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
MCCULLOCH, MICHAEL
TURNER, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howrey & Simon

STREET: 1299 Pennsylvania Avenue N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/466,422

FILING DATE: 17-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/037,751

Filing DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Halluin, Albert P

REGISTRATION NUMBER: 25,277

REFERENCE/DOCKET NUMBER: 00801.0140.999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-463-8109

TELEFAX: 650-463-8400

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 6446 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: Genomic RNA

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-466-422-5

Alignment Scores:

Pred. No.:

48.4

Length:

1362

Pred. No.: 17.2 Length: 6446
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x US-09-466-422-5 (1-6446)

QY 4 LysTpProTtpTtpProTtpArg 12
DB 6213 AAGUGCCUUGGUGGCGGCGCA 6239

RESULT 9

US-09-252-991A-9269
Sequence 9269, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 9269

LENGTH: 1278

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9269

Alignment Scores:

Pred. No.: 45.3 Length: 1278
Score: 63.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.23% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x US-09-252-991A-9269 (1-1278)

QY 5 TtpProTtpTtpProTtpArg 11

DB 885 TGGCCCTGGTGGCCCTGGCGA 905

RESULT 10

US-09-252-991A-9298/c

Sequence 9298, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 9298

LENGTH: 1362

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9298

Alignment Scores:

Pred. No.:

48.4

Length:

1362

Score: 63.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.23% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-35 (1-13) x US-09-252-991A-9298 (1-1362)

QY 5 TtpProTtpTtpProTtpArg 11
Db 623 TGGCCCTGGTGGCCCTGGCGA 603

RESULT 11

US-09-252-991A-9281
; Sequence 9281, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9281
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9281

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
54.3	1521	63.00	7
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		69.23%	Indels: 0
DB:		4	Gaps: 0

US-09-444-281-35 (1-13) x US-09-252-991A-9281 (1-1521)

QY 5 TtpProTtpTtpProTtpArg 11
Db 992 TGGCCCTGGTGGCCCTGGCGA 1012

RESULT 12

US-08-159-784-1/C
; Sequence 1, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4031
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-1

Alignment Scores:	Pred. No.:	Length:	Score:	Matches:
	332	4031	60.00	7
Percent Similarity:		77.78%	Conservative:	0
Best Local Similarity:		77.78%	Mismatches:	2
Query Match:		65.93%	Indels:	0
DB:		1	Gaps:	0

US-09-444-281-35 (1-13) x US-08-159-784-1 (1-4031)

QY 3 LysLysTtpProTtpTtpProTtpArg 11
Db 1671 AAATCCTGGTCTGGTGGCCCTGGAGG 1645

RESULT 13

US-09-313-294A-5936/C
; Sequence 5936, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5936
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350926H1
; NAME/KEY: unsure
; LOCATION: 2, 35, 208, 242, 276, 279
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5936

Alignment Scores:	Pred. No.:	Length:	Score:	Matches:
	36.7	292	58.00	6
Percent Similarity:		100.00%	Conservative:	0
Best Local Similarity:		100.00%	Mismatches:	0
Query Match:		63.74%	Indels:	0
DB:		4	Gaps:	0

US-09-444-281-35 (1-13) x US-09-313-294A-5936 (1-292)

QY 5 TtpProTtpTtpProTtp 10
Db 101 TGGCCCTGGTGGCCCTGGG 84

RESULT 14

US-08-481-337A-1
; Sequence 1, Application US/08481337A
; Patent No. 5863738
; GENERAL INFORMATION:
; APPLICANT: TEN DIJKE, Peter
; APPLICANT: HELDIN, Carl-Henrik
; APPLICANT: MIYAZONO, Kohei
; APPLICANT: SAMPATH, Kuber T.
; TITLE OF INVENTION: Morphogenic Protein-Specific Cell
; TITLE OF INVENTION: Surface Receptors and Uses Therefor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.337A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-097CF2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1509 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1509
; OTHER INFORMATION: /product= "Human ALK1"
US-08-481-337A-1
Alignment Scores:
Pred. No.: 203 Length: 1509
Score: 58.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.74% Indels: 0
DB: 2 Gaps: 0
US-09-444-281-35 (1-13) x US-08-481-337A-1 (1-1509)
QY 5 TtpProTtpTtpProTtp 10
Db 389 TGGCCCTGGTGGCCCTGG 406
RESULT 15
US-08-696-268B-1
; Sequence 1, Application US/08696268B
; Patent No. 5968752
; GENERAL INFORMATION:
; APPLICANT: ICHIO, Hidenori
; APPLICANT: NISHITOH, Hideki
; APPLICANT: SAMPATH, Kuber T.
; TITLE OF INVENTION: NOVEL SIGNALLING RECEPTOR FOR
; TITLE OF INVENTION: MORPHOGENIC PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696.268B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1509 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1509
; OTHER INFORMATION: /product= "Human ALK-1"
US-08-696-268B-1
Alignment Scores:
Pred. No.: 203 Length: 1509
Score: 58.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 63.74% Indels: 0
DB: 2 Gaps: 0
US-09-444-281-35 (1-13) x US-08-696-268B-1 (1-1509)
QY 5 TtpProTtpTtpProTtp 10
Db 389 TGGCCCTGGTGGCCCTGG 406
Search completed: May 11, 2004, 16:58:17
Job time : 51.92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:35:33 ; Search time 281.32 Seconds
(without alignments)
209.334 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKWFWPWRK 13

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US09444281/runat_07052004_171139_2611/app_query.fasta_1.398
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09444281@cgn_1_271_@runat_07052004_171139_2611
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq2.*
19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

1	76	83.5	72	16	US-10-395-896-32	Sequence 32, Appl
2	76	83.5	139	16	US-10-395-896-31	Sequence 31, Appl
3	75	82.4	121	16	US-10-395-896-28	Sequence 28, Appl
4	75	82.4	121	16	US-10-395-896-28	Sequence 28, Appl
5	75	82.4	123	16	US-10-395-896-29	Sequence 29, Appl
6	75	82.4	166	16	US-10-395-896-30	Sequence 30, Appl
7	73	80.2	68	9	US-09-909-652-6	Sequence 6, Appl
8	73	80.2	550	15	US-10-076-816-60	Sequence 60, Appl
9	73	80.2	6446	10	US-09-962-527-5	Sequence 5, Appl
10	70	76.9	114	16	US-10-395-896-64	Sequence 64, Appl
11	70	76.9	114	16	US-10-395-896-65	Sequence 65, Appl
12	66	72.5	207	15	US-10-252-773-25	Sequence 25, Appl
13	64	70.3	39	15	US-10-252-773-15	Sequence 15, Appl
14	64	70.3	53	15	US-10-252-773-17	Sequence 17, Appl
15	64	70.3	54	15	US-10-252-773-19	Sequence 19, Appl
16	64	70.3	557	13	US-10-027-632-53929	Sequence 53929, A
17	64	70.3	557	13	US-10-027-632-321717	Sequence 321717, A
18	64	70.3	557	16	US-10-027-632-53929	Sequence 53929, A
19	64	70.3	557	16	US-10-027-632-321717	Sequence 321717, A
20	63	69.2	122	9	US-09-783-590-1444	Sequence 1444, Ap
21	63	69.2	1437	9	US-09-815-242-7766	Sequence 7766, Ap
22	63	69.2	1908	9	US-09-815-242-7716	Sequence 7716, Ap
23	63	69.2	1908	13	US-10-282-122A-39517	Sequence 39517, A
24	61.5	67.6	571	13	US-10-425-114-14181	Sequence 14181, A
25	60	65.9	577	15	US-10-029-386-12845	Sequence 12845, A
26	60	65.9	1526	13	US-10-424-599-17644	Sequence 17644, A
27	60	65.9	2403	15	US-10-156-761-4600	Sequence 4600, Ap
28	60	65.9	14739	15	US-10-017-161-1557	Sequence 1557, Ap
29	60	65.9	14739	16	US-10-292-798-1245	Sequence 1245, Ap
30	60	65.9	34570	12	US-09-997-722-157	Sequence 157, App
31	60	65.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
32	59.5	65.4	1224	16	US-10-369-493-40678	Sequence 40678, A
33	59	64.8	502	13	US-10-027-632-44536	Sequence 44536, A
34	59	64.8	502	16	US-10-027-632-44536	Sequence 44536, A
35	59	64.8	627	13	US-10-027-632-94574	Sequence 94574, A
36	59	64.8	627	16	US-10-027-632-94574	Sequence 94574, A
37	59	64.8	669	16	US-10-260-238-659	Sequence 659, App
38	59	64.8	837	13	US-10-424-599-17642	Sequence 17642, A
39	59	64.8	1825	16	US-10-116-275-321	Sequence 321, App
40	59	64.8	2218	10	US-09-820-790-1	Sequence 1, Appli
41	58	63.7	228	9	US-09-933-876-3798	Sequence 3798, Ap
42	58	63.7	228	11	US-09-923-876-3798	Sequence 3798, Ap
43	58	63.7	429	13	US-10-424-599-1131	Sequence 1131, Ap
44	58	63.7	496	13	US-10-424-599-92875	Sequence 92875, A
45	58	63.7	767	13	US-10-027-632-132877	Sequence 132877, A

ALIGNMENTS

RESULT 1
US-10-395-896-32
; Sequence 32, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; TITLE OF INVENTION: ANTI-INFECTIVE PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Exemplary "template"
; US-10-395-896-28

Alignment Scores:
Pred. No.: 0.129 Length: 121
Score: 75.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best local Similarity: 90.00% Mismatches: 0
Query Match: 82.42% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-35 (1-13) x US-10-395-896-28 (1-121)

QY 4 LysTrrProTrrPrrProTrrArgArgLys 13
   ::|||::|||::|||::|||::|||::|||
DB 54 CGTTGCCCGTGGTGGCCGTGGCGTGCATA 83

RESULT 4
US-10-395-896-62/c
; Sequence 62, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; TITLE OF INVENTION: ANTI-INFECTIVE PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid template encoding anti-infective
; OTHER INFORMATION: peptide 11B7 and anionic spacer peptide S11 or S12
; US-10-395-896-62

Alignment Scores:
Pred. No.: 0.129 Length: 121
Score: 75.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best local Similarity: 90.00% Mismatches: 0
Query Match: 82.42% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-35 (1-13) x US-10-395-896-62 (1-121)

QY 4 LysTrrProTrrPrrProTrrArgArgLys 13
   ::|||::|||::|||::|||::|||::|||
DB 68 CGTTGCCCGTGGTGGCCGTGGCGTGCATA 99

RESULT 5
US-10-395-896-29
; Sequence 29, Application US/10395896
; Publication No. US20030219854A1

```


Alignment Scores:
Pred. No.: 0.953 Length: 550
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-35 (1-13) x US-10-076-816-60 (1-550)

QY 4 LysTrpProTrrpProTrrpArgArg 12
Db 415 AAATGGCAATGTCGGCTTGGCCAGA 441

RESULT 9

; Sequence 5, Application US/0962527
; Publication No. US20030049813A1
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; HOLTZ, R. BARRY
; MCCULLOCH, MICHAEL
; TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,527
; FILING DATE: 24-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: 10-march-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-962-527-5
Alignment Scores:
Pred. No.: 9.69 Length: 6446
Score: 73.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.22% Indels: 0
DB: 10 Gaps: 0

US-09-444-281-35 (1-13) x US-09-962-527-5 (1-6446)

QY 4 LysTrpProTrrpProTrrpArgArg 12
Db 6213 AAUGGCGCCUUGGUGGCGGCGCGA 6239

RESULT 10

; Sequence 64, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid template encoding precursor peptide
; OTHER INFORMATION: 11B25 and anionic spacer peptide S21.
US-10-395-896-64

Alignment Scores:
Pred. No.: 0.513 Length: 114
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 76.92% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-35 (1-13) x US-10-395-896-64 (1-114)

QY 4 LysTrpProTrrpProTrrpArgArg 12
Db 36 CGTTGGCCGTCGTGGCGTCGC 62

RESULT 11

; Sequence 65, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid template encoding precursor peptide

; OTHER INFORMATION: 11B25 and anionic spacer peptide S21.
US-10-395-896-65

Alignment Scores: 114
Pred. No.: 0.513 Length: 114
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 76.92% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-35 (1-13) x US-10-395-896-65 (1-114)

QY 4 LysTyrProTyrProTyrProTyrArg 12

Db 79 CGTTGGCGGTGGTGGCGGTGGCGTGC 53

RESULT 12

US-10-252-773-25
; Sequence 25, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA PCRIL
; OTHER INFORMATION: construct
US-10-252-773-25

Alignment Scores: 207
Pred. No.: 2.85 Length: 207
Score: 66.00 Matches: 6
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 0
Query Match: 72.53% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-35 (1-13) x US-10-252-773-25 (1-207)

QY 2 LeuLysTyrProTyrProTyrProTyrArg 11

Db 163 ATTAGGAGTGGCTTGGTGGCTTGGAAA 192

RESULT 13

US-10-252-773-15
; Sequence 15, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773

; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Oligonucleotide
US-10-252-773-15

Alignment Scores: 39
Pred. No.: 1.05 Length: 39
Score: 64.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 70.33% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-35 (1-13) x US-10-252-773-15 (1-39)

QY 3 LysLysTyrProTyrProTyrArg 11

Db 1 AGGAGTGGCTTGGTGGCTTGGAAA 27

RESULT 14

US-10-252-773-17/c
; Sequence 17, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 53
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-252-773-17

Alignment Scores: 53
Pred. No.: 1.4 Length: 53
Score: 64.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 70.33% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-35 (1-13) x US-10-252-773-17 (1-53)

QY 3 LysLysTyrProTyrProTyrArg 11

Db 53 AGGAGTGGCTTGGTGGCTTGGAAA 27

RESULT 15

```

US-10-252-773-19/c
; Sequence 19, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 54
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Oligonucleotide
US-10-252-773-19

```

```

Alignment Scores:
Pred. No.: 1.43 Length: 54
Score: 64.00 Matches: 6
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 70.33% Indels: 0
DB: 15 Gaps: 0

```

US-09-444-281-35 (1-13) x US-10-252-773-19 (1-54)

```

QY 3 LysLysTrpProTrpProTrpArg 11
Db 54 AGAGATGGCCTTGGTGGCTTGGAAA 28

```

Search completed: May 11, 2004, 17:07:32
Job time : 286.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:00:48 ; Search time 1721.2 Seconds
(without alignments)
225.545 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKXPPWPRRK 13

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09444281/runat_07052004_171138_2579/app_query.fasta_1.398
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -IOOPCI=0 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281 -CGN 1 1 3596 @runat_07052004_171138_2579 -NCEU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mus.*
23: em_gss_mam.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	81.3	522	14	CF479395 RTW3_23
2	70	76.9	263	12	BI976779 485902 MA
3	69	75.8	330	10	BE024584 894003H02
4	69	75.8	390	28	AZ518927 RPCI-11-6
5	69	75.8	757	29	AG185290 Pan ttrgl
6	68	74.7	349	12	BJ476852 BJ476852
7	68	74.7	369	12	BJ474341 BJ474341
8	68	74.7	381	12	BJ471810 BJ471810
9	68	74.7	411	12	BJ473016 BJ473016
10	68	74.7	415	10	BE237369 146629 MA
11	68	74.7	420	9	AV933841 AV933841
12	68	74.7	440	12	BJ471261 BJ471261
13	68	74.7	446	9	AU089922 AU089922
14	68	74.7	446	9	AU198144 AU198144
15	68	74.7	448	9	AU198162 AU198162
16	68	74.7	464	9	AV935002 AV935002
17	68	74.7	471	12	BJ475478 BJ475478
18	68	74.7	472	12	BJ475115 BJ475115
19	68	74.7	472	12	BJ476710 BJ476710
20	68	74.7	473	9	AV937233 AV937233
21	68	74.7	474	12	BJ473880 BJ473880
22	68	74.7	475	9	AU089934 AU089934
23	68	74.7	489	12	BJ477184 BJ477184
24	68	74.7	500	12	BJ475153 BJ475153
25	68	74.7	501	12	BJ470669 BJ470669
26	68	74.7	504	12	BU475306 BU475306
27	68	74.7	506	12	BJ471761 BJ471761
28	68	74.7	506	12	BJ473882 BJ473882
29	68	74.7	509	12	BJ471258 BJ471258
30	68	74.7	515	12	BJ469760 BJ469760
31	68	74.7	530	12	BU472623 BU472623
32	68	74.7	531	12	BU475836 BU475836
33	68	74.7	532	9	AV932159 AV932159
34	68	74.7	539	14	CF326637 JMT1--06-
35	68	74.7	550	14	CF326444 JMT1--06-
36	68	74.7	559	14	CF325873 JMT1--04-
37	68	74.7	578	9	AU082117 AU082117
38	68	74.7	595	14	CF326279 JMT1--05-
39	68	74.7	661	14	CB681162 OSJNEF07B
40	67	73.6	1172	28	BZ552545 pacs1-60
41	66	72.5	352	9	AU198258 AU198258
42	66	72.5	396	9	AL916683 AL916683
43	66	72.5	434	10	BE453064 BE453064
44	66	72.5	445	12	BG305131 BG305131
45	66	72.5	568	13	CA155445 SCACRZ310

ALIGNMENTS

RESULT 1
CF479395 522 bp mRNA linear EST 08-SEP-2003
LOCUS RTW3_23 A01.G1 A022 Well-watered loblolly pine roots WW3 Pinus
DEFINITION taeda_cDNA clone RTW3_23 A01_A022 5', mRNA sequence.
ACCESSION CF479395.1 GI:34508264
VERSION Pinus taeda (loblolly pine)
KEYWORDS EST.
SOURCE Pinus taeda
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 522)

AUTHORS

Pratt, L., Cordomier-Pratt, M.-M., Lorenz, W.W., Dean, J.,
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., and
Neale, D.

TITLE
JOURNAL

An EST database from well-watered loblolly pine (*Pinus taeda*) roots
Unpublished (2003)

COMMENT

Other ESTs: RTW3 23 A01.b1 A022
Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of
Forestry, University of Georgia; plant material prepared at the
University of Florida; sequencing done in the Laboratory for
Genomics and Bioinformatics, University of Georgia. Sequence ends
have been trimmed to exclude vector and regions below phred quality
16. Three-prime sequences are presented as their reverse complement
and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES

source

1. 522
Location/Qualifiers
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="CCLONES"
/db_xref="taxon:3352"
/clone="RTW3.23.A01.A022"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Well-watered loblolly pine roots W3"
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from polyA+ RNA from loblolly pine
(Pinus taeda) roots watered to pot capacity every other
day. Pre-dawn water potential remained -0.3 MPa +/-0.1.
Roots were harvested for RNA isolation. Double-stranded
cDNA was cloned unidirectionally into pSL1180. Inserts
excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Alignment Scores: 172 Length: 522
Pred. No.: 74.00 Matches: 8
Score: 100.00% Conservative: 3
Percent Similarity: 72.73% Mismatches: 0
Best Local Similarity: 81.32% Indels: 0
Query Match: 14 Gaps: 0
DB:

US-09-444-281-35 (1-13) x CF479395 (1-522)

QY 3 LysLysTrpProTrpTrpProTrpArgArglys 13
DB 284 AGGAGTGGCCCTGTGTGCTGGCGGGGAGG 316

RESULT 2

B1976779

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

source

1 (bases 1 to 263)

Smith, T.P.I., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,

Bennett, G.L., Heaton, M.P., Lagreid, W.W., Rohrer, G.A.,

Chitko-McKown, C.G., Perlea, G., Holt, I., Karamycheva, S., Liang, F.,

Quackenbush, J., and Keefe, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 114 row: F column: 23

Seq primer: ATTAGTGCACACTATAG.

Location/Qualifiers

1. 263

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 2BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from testis, thymus,

semitendinosus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

ORIGIN

Alignment Scores: 201 Length: 263
Pred. No.: 70.00 Matches: 8
Score: 100.00% Conservative: 1
Percent Similarity: 88.89% Mismatches: 0
Best Local Similarity: 76.92% Indels: 0
Query Match: 12 Gaps: 0
DB:

US-09-444-281-35 (1-13) x B1976779 (1-263)

QY 4 LysTrpProTrpTrpProTrpArgArg 12

DB 126 AAATGGCCATGGTGGCTTGGCGCAAA 152

RESULT 3

BE024584

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 330)

Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadales; Chlamydomonas.

Unpublished (2000)

Contact: Elizabeth H. Harris

Duke University

Durham, NC 27708-1000, USA

Tel: 919 613 8164

Fax: 919 613 8177

Email: chlamy@duke.edu.

Location/Qualifiers

1. 330

source

BE024584 330 bp mRNA linear EST 06-JUN-2000
894003H02.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BE024584

VERSION BE024584.1

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii

ORGANISM Chlamydomonas reinhardtii

REFERENCE 1 (bases 1 to 330)

AUTHORS Grosse, W.M., Harris, E., Lefebvre, P.,

McDermott, J.P., Silflow, C., Stern, D., and Surzycki, R.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants; project phase 2

Unpublished (2000)

Contact: Elizabeth H. Harris

Duke University

Durham, NC 27708-1000, USA

Tel: 919 613 8164

Fax: 919 613 8177

Email: chlamy@duke.edu.

Location/Qualifiers

1. 330

source

```

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, lambda Zap
II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
polyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 335 Length: 330
Score: 69.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 75.82% Indels: 0
DB: 10 Gaps: 0

```

US-09-444-281-35 (1-13) x BE024584 (1-330)

QY 2 LeuLysLysTrpProTTPProTTPArgArg 12

DB 66 CTAGACGGTGGCGGTGGTGGCGGTGGCGCG 98

RESULT 4

AZ518927/c

```

LOCUS RP43-059F08.TJ RP43-059F08.TJ Homo sapiens genomic clone RP43-11-67B18,
DEFINITION genomic survey sequence.

```

ACCESSION AZ518927

VERSION AZ518927.1

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other GSSs: RP43-059F08.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the human BAC library RP43-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genet cs (info@resgen.com). BAC end search page:

http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.

This BAC end was generated during the R&D process and may have

higher chance of clone tracking errors.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..390

/organism="Homo sapiens"

FEATURES

source

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

```

/mol_type="genomic DNA"
/db_xref="GDB:7525385"
/db_xref="taxon:9606"
/clone="RP43-11-67B18"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RP43-11"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
RP43-11 Human Male BAC Library"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 408 Length: 390
Score: 69.00 Matches: 10
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 75.82% Indels: 0
DB: 28 Gaps: 0

```

US-09-444-281-35 (1-13) x AZ518927 (1-390)

QY 2 LeuLysLysTrpProTTPProTTPArgArgLys 13

DB 189 TTACAAAATCCCCCTGGTGGCCCTGGAGGAGGAG 154

RESULT 5

AG185290/c

LOCUS AG185290

DEFINITION Pan troglodytes DNA, clone: RP43-059F08.TJ, genomic survey

sequence.

ACCESSION AG185290

VERSION AG185290.1

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan 305-8565, Japan

(E-mail: chimpes@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/)

Tel: 81-45-503-9111, Fax: 81-45-503-9170

Clones are derived from the chimpanzee BAC library RP43-43 This BAC

end was generated during the R&D process and may have higher chance

of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

Location/Qualifiers

1..757

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="RP43-059F08.TJ"

/sex="male"

/cell_type="Lymphocytes"

/clone_lib="RP43-43 Chimpanzee Male BAC Library"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Score: 69.00 Matches: 10
 Percent Similarity: 91.67% Conservative: 1
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 75.82% Indels: 0
 DB: 29 Gaps: 0

US-09-444-281-35 (1-13) x AG185290 (1-757)

QY 2 LeuylsytTtpProTtpTtpArgArg 13
 DB 279 TTACAAAATCCCCCTGGTGGCCCTGGAGGAG 244

RESULT 6

BJ476852 349 bp mRNA linear EST 23-MAY-2002
 LOCUS BJ476852 K. Sato unpublished cDNA library, cv. Haruna Nijo adult.
 DEFINITION heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
 clone baal37g15 3', mRNA sequence.

ACCESSION BJ476852
 VERSION BJ476852.1 GI:21155348

KEYWORDS

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 349)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source
 1..349
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Haruna Nijo"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="baal37g15"
 /tissue_type="top three leaves"
 /dev_stage="adult, heading stage"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:
 Pred. No.: 455 Length: 349
 Score: 68.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 74.73% Indels: 0
 DB: 12 Gaps: 0

US-09-444-281-35 (1-13) x BJ476852 (1-349)

QY 5 TtpProTtpTtpProTtpArgArg 12
 DB 275 TGGCCGTGGTGGCCGTGGCGGCA 298

RESULT 7

BJ474341 369 bp mRNA linear EST 23-MAY-2002
 LOCUS BJ474341 K. Sato unpublished cDNA library, cv. Haruna Nijo adult.
 DEFINITION heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
 clone baal3el8 3', mRNA sequence.

ACCESSION BJ474341

VERSION BJ474341.1 GI:21152844

KEYWORDS

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 369)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source
 1..369
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Haruna Nijo"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="baal3el8"
 /tissue_type="top three leaves"
 /dev_stage="adult, heading stage"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:
 Pred. No.: 486 Length: 369
 Score: 68.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 74.73% Indels: 0
 DB: 12 Gaps: 0

US-09-444-281-35 (1-13) x BJ474341 (1-369)

QY 5 TtpProTtpTtpProTtpArgArg 12
 DB 277 TGGCCGTGGTGGCCGTGGCGGCA 300

RESULT 8

BJ471810 381 bp mRNA linear EST 23-MAY-2002
 LOCUS BJ471810 K. Sato unpublished cDNA library, cv. Haruna Nijo adult.
 DEFINITION heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
 clone baal21m20 5', mRNA sequence.

ACCESSION BJ471810

VERSION BJ471810.1 GI:21150313

KEYWORDS

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 381)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source
 1..381
 /organism="Hordeum vulgare subsp. vulgare"

```

/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baal21m20"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo adult, heading stage top three leaves"

ORIGIN
Alignment Scores:
Pred. No.: 505 Length: 381
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 12 Gaps: 0

US-09-444-281-35 (1-13) x BJ471810 (1-381)

QY 5 TtpProTtpTtpProTtpArgArg 12
Db 183 TGGCGGTGGTGGCGGTGGCGGCA 160

RESULT 9
BJ473016/c
LOCUS BJ473016 411 bp mRNA linear EST 23-MAY-2002
DEFINITION
heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
clone baal40h02 5', mRNA sequence.
ACCESSION BJ473016
VERSION BJ473016.1 GI:21151519
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 411)
Sato, K., Saisho, P. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..411
Location/Qualifiers
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baal40h02"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo adult, heading stage top three leaves"

ORIGIN
Alignment Scores:
Pred. No.: 552 Length: 411
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 12 Gaps: 0

US-09-444-281-35 (1-13) x BJ473016 (1-411)

```

```

QY 5 TtpProTtpTtpProTtpArgArg 12
Db 177 TGGCGGTGGTGGCGGTGGCGGCA 154

RESULT 10
BE237369
LOCUS BE237369 415 bp mRNA linear EST 25-APR-2001
DEFINITION
146629 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE237369
VERSION BE237369.1 GI:9022087
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 415)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F.,
Quackenbush, J. and Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGAG
Plate: 47 row: J column: 10
Seq primer: ATTAGTGCACCTATAG.
Location/Qualifiers
1..415
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4BOV"
/notes="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."

ORIGIN
Alignment Scores:
Pred. No.: 559 Length: 415
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 10 Gaps: 0

US-09-444-281-35 (1-13) x BE237369 (1-415)

QY 4 LysTtpProTtpTtpProTtpArg 11
Db 280 AAATGCCCATGGTGGCTTGGCGC 303

RESULT 11
AV933841/c
LOCUS AV933841 420 bp mRNA linear EST 18-JAN-2002
DEFINITION
AV933841 K. Sato unpublished cDNA library, cv. Haruna Nijo adult,

```


heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
clone baal9005 5', mRNA sequence.

ACCESSION AV933841
VERSION AV933841.1 GI:18229638

KEYWORDS EST.

ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

FEATURES
source 1..420
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baal9005"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo adult, heading stage top three leaves"

Alignment Scores:

Pred. No.: 567 Length: 420
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AV933841 (1-420)

QY 5 TtpProTtpTrpProTtpArg 12
|||||
DB 186 TGCCGCTGGTGGCGTGGCGGCGA 163

RESULT 12

LOCUS B471261/c

DEFINITION B471261 440 bp mRNA linear EST 23-MAY-2002
heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
clone baal17107 5', mRNA sequence.

ACCESSION B471261

VERSION B471261.1 GI:21149764

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 440)

AUTHORS Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

FEATURES
source 1..440
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baal17107"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 599 Length: 440
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 12 Gaps: 0

US-09-444-281-35 (1-13) x B471261 (1-440)

QY 5 TtpProTtpTrpProTtpArg 12
|||||

DB 206 TGCCGCTGGTGGCGTGGCGGCGA 183

RESULT 13

LOCUS AU089922/c

DEFINITION AU089922 446 bp mRNA linear EST 19-APR-2000
heading stage Hordeum vulgare subsp. vulgare Upper three leaves at
haruna lib1_121, mRNA sequence.

ACCESSION AU089922

VERSION AU089922.1 GI:7613350

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 446)

AUTHORS Sato, K., Takanashi, H. and Takeda, K.
TITLE Hordeum vulgare subsp. vulgare cDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Kazuhiro Sato

Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazsato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/.

Location/Qualifiers

FEATURES
source 1..446
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="haruna lib1_121"
/tissue_type="Upper three leaves at heading stage"
/clone_lib="Hordeum vulgare subsp. vulgare Upper three
leaves at heading stage"

ORIGIN

Alignment Scores:

Pred. No.: 608 Length: 446
Score: 68.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.73% Indels: 0
DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AU089922 (1-446)

QY 5 TtpProTtpTrpProTtpArgArg 12
 Db 188 TGGCGCTGGTGGCGCTGGCGCGA 165

RESULT 14

AUI98144 446 bp mRNA linear EST 03-APR-2002
 AUI98144/c
 LOCUS AUI98144 Rice green shoot Oryza sativa (japonica cultivar-group)
 DEFINITION CDNA clone S15951, mRNA sequence.

ACCESSION AUI98144

VERSION AUI98144.1 GI:14714211

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 446)

AUTHORS Sasaki,T. and Yamamoto,K.

TITLE Rice cDNA from green shoot (2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Takuji Sasaki

National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/

PROJECT = 'RGP'.

FEATURES

source
 1..446
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S15951"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

ORIGIN

Alignment Scores:
 Pred. No.: 608 Length: 446
 Score: 68.00 Matches: 7
 Percent Similarity: 90.91% Conservative: 3
 Best Local Similarity: 63.64% Mismatches: 1
 Query Match: 74.73% Indels: 0
 DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AUI98144 (1-446)

QY 3 LysLysTtpProTtpTrpProTtpArgArgLys 13

Db 351 CGCGCTGGCTGGTGGCGCTGGCGCGG 319

RESULT 15

AUI98162/c
 LOCUS AUI98162 448 bp mRNA linear EST 03-APR-2002
 DEFINITION AUI98162 Rice green shoot Oryza sativa (japonica cultivar-group)
 CDNA clone S16019, mRNA sequence.

ACCESSION AUI98162

VERSION AUI98162.1 GI:14714231

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 448)

AUTHORS Sasaki,T. and Yamamoto,K.

TITLE Rice cDNA from green shoot (2001)

JOURNAL COMMENT

Unpublished (2001)
 Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 PROJECT = 'RGP'.

FEATURES

source
 1..448
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S16019"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

ORIGIN

Alignment Scores:
 Pred. No.: 612 Length: 448
 Score: 68.00 Matches: 7
 Percent Similarity: 90.91% Conservative: 3
 Best Local Similarity: 63.64% Mismatches: 1
 Query Match: 74.73% Indels: 0
 DB: 9 Gaps: 0

US-09-444-281-35 (1-13) x AUI98162 (1-448)

QY 3 LysLysTtpProTtpTrpProTtpArgArgLys 13

Db 314 CGCGCTGGCTGGTGGCGCTGGCGCGG 282

Search completed: May 11, 2004, 16:56:32

Job time : 1725.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 13:53:18 ; Search time 2221.92 Seconds
(without alignments)
234.084 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPWPWRK 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09444281/runat_07052004_171138_2568/app.query.fasta_1.398
-DB=GenEmbl -QFW=fastap -SUFFIX=zge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281@cgn_1_3608 @runat_07052004_171138_2568 -NCPUS=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -KONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	86	100.0	108	6	BD247523	BD247523 Method fo	BD247529 Method fo BD247520 Method fo BD247517 Method fo AX357080 Sequence BD128612 Method fo AR404925 Sequence AR226388 Sequence AR282754 Sequence X67340 Bos taurus AR173324 Sequence AX098418 Sequence AL672300 Mouse DNA AC113900 Rattus no AC095460 Rattus no U08210 Mus musculu BC051649 Mus muscu AF289665 Mus muscu AP004382 Oryza sat AC091250 Mus muscu AJ300832 Delftia a Continuation (8 of BD273664 Peptides AF210429 Mus muscu BC028879 Mus muscu BD095993 Immunocass BD103482 Immunocass AF166097 Mus muscu AC132933 Mus muscu AC120860 Mus muscu AC139212 Mus muscu AC102491 Mus muscu AY354516 Trypanoso AC104488 Trypanoso BD273655 Peptides BD273657 Peptides BD273659 Peptides AF005217 Corynebact AX653694 Sequence X17342 Arabidopsis 109762 Arabidopsis AC005724 Arabidops AP003141 Oryza sat BC029697 Mus muscu Z22923 M.musculus
2	86	100.0	114	6	BD247529		
3	86	100.0	151	6	BD247520		
4	75	87.2	114	6	BD247517		
5	74	86.0	68	6	AX357080		
6	70	81.4	39	6	BD128612		
7	70	81.4	69	6	AR404925		
8	70	81.4	211	6	AR226388		
9	70	81.4	211	6	AR282754		
10	70	81.4	550	4	BTINDLCD		
11	70	81.4	6446	6	AR173324		
12	70	81.4	6446	6	AX098418		
13	70	81.4	149991	10	AL672300		
14	70	81.4	195764	2	AC113900		
15	70	81.4	255121	2	AC095460		
16	69	80.2	2651	10	MMU08210		
17	69	80.2	3643	10	BC051649		
18	69	80.2	107257	10	AF289665		
19	69	80.2	141040	8	AP004382		
20	69	80.2	201395	10	AC091250		
21	68	79.1	7505	1	CAC300832		
22	68	79.1	110000	2	BX255276_07		
23	67	77.9	207	6	BD273664		
24	67	77.9	456	10	AF210429		
25	67	77.9	1011	10	BC028879		
26	67	77.9	1040	6	BD095993		
27	67	77.9	1040	6	BD103482		
28	67	77.9	1040	10	AF166097		
29	67	77.9	180007	2	AC132933		
30	67	77.9	185784	2	AC120860		
31	67	77.9	203000	2	AC139212		
32	67	77.9	204981	2	AC102491		
33	66	76.7	2442	3	AX354516		
34	66	76.7	138208	2	AC104488		
35	65	75.6	39	6	BD273655		
36	65	75.6	53	6	BD273657		
37	65	75.6	54	6	BD273659		
38	65	75.6	300750	1	AF005217		
39	64.5	75.0	1539	6	AX653694		
40	64.5	75.0	3850	8	ATPHYB		
41	64.5	75.0	6509	8	ATHPTOCHB		
42	64.5	75.0	88413	8	AC005724		
43	64.5	75.0	141275	8	AP003141		
44	64	74.4	2718	10	BC029697		
45	64	74.4	19479	10	MMA2IXCOA		

ALIGNMENTS

RESULT 1

PC	C12N15/09, C07K1/12, C07K1/18, C12N1/15, C12N1/19, C12N1/21, C12P21/02
PC	02
PC	(C12P21/02, C12R1/19), C12N15/00
CC	Synthesized oligonucleotide used as a template for PCR FH
Key	Location/Qualifiers
FT	source 1..114
FT	/organism='Artificial Sequence'. Location/Qualifiers
FEATURES	1..114
source	/organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0286 Length: 114
Score:	86.00 Matches: 12
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	6 Gaps: 0
US-09-444-281-36 (1-12) x BD247529 (1-114)	
Qy	1 lleLeuArgTrpProTrpTrpProTrpArgArgLys 12
Db	47 ATTCTGCGTTGGCCGTTGGTGGCGCGTGGCGTGCAGAA 82

BD247520	BD247520	151 bp	DNA	linear	PAT 17-JUN-2002
LOCUS	Method for effectively producing antibacterial cationic peptides in host cells.				
DEFINITION	BD247520				
ACCESSION	BD247520.1 GI:33057290				
VERSION	JP 2002530114-A/14.				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 151)				
AUTHORS	Burian, J. and Bartfeld, D.				
TITLE	Method for effectively producing antibacterial cationic peptides in host cells				
JOURNAL	Patent: JP 2002530114-A 14 17-SEP-2002;				
COMMENT	MICROLOGIX BIOTECH INC				
	OS Artificial Sequence				
	PN JP 2002530114-A/14				
	PD 17-SEP-2002				

PI	JAN BURIAN, BARTIAN, FELD
PC	C12N15/09, C07K1/18, C12N1/15, C12N1/19, C12N1/21, C12P21/02
PC	//
PC	(C12P21/02, C12R1/19), C12N15/00
CC	Synthesized oligonucleotide used as a template for PCR FH
Key	Location/Qualifiers
FT	source 1..151
FT	/organism='Artificial Sequence'.
FEATURES	Location/Qualifiers
source	1..151
	/organism="synthetic construct"
	/mol_type="genomic DNA"
	/db_xref="taxon:32630"
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0382
Score:	Length: 151
	Matches: 12
Percent Similarity:	Conservative: 0
Best Local Similarity:	Mismatches: 0
Query Match:	Indels: 0
DR:	Gaps: 0

US-09-444-281-36 (1-12) x BD247520 (1-151)

QY 1 IleleuArgTrpProTrpTrpProTrpArgArgLys 12
 Db 41 ATTCTGGTGGCGGTGGTGGCGGTGGCGTAA 76

RESULT 4
 BD247517 114 bp DNA linear PAT 17-JUL-2003
 LOCUS Method for effectively producing antibacterial cationic peptides in
 DEFINITION host cells.
 ACCESSION BD247517
 VERSION BD247517.1 GI:33057287
 KEYWORDS JP 2002530114-A/11.
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 114)
 AUTHORS Burian, J. and Bartfeld, D.
 TITLE Method for effectively producing antibacterial cationic peptides in
 JOURNAL host cells
 PATENT: JP 2002530114-A 11 17-SEP-2002;
 COMMENT MICROLOGIX BIOTECH INC
 OS Artificial Sequence
 PN JP 2002530114-A/11
 PD 17-SEP-2002
 PF 19-NOV-1999 JP 2000584088
 PR 20-NOV-1998 US 60/109218
 PI JAN BURIAN, DANIEL BARTFELD
 PC C12N15/09, C07K1/12, C07K1/18, C12N1/15, C12N1/19, C12N1/21, C12P21/
 PC 02//
 PC C12P21/02, C12P1/19, C12N15/00
 CC Synthesized oligonucleotide used as a template for PCR FH

KEY Key
 FT source
 FT Location/Qualifiers
 1. 114
 /organism="Artificial Sequence".
 Location/Qualifiers
 1. 114
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

FEATURES
 source

ORIGIN

Alignment Scores:
 Pred. No.: 0.532 Length: 114
 Score: 75.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 90.00% Mismatches: 0
 Query Match: 87.21% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x BD247517 (1-114)

QY 3 ArgTrpProTrpTrpProTrpArgArgLys 12
 Db 50 AAATGGCGGTGGTGGCGGTGGCGGTAA 79

RESULT 5
 AX357080 68 bp DNA linear PAT 13-FEB-2002
 LOCUS Sequence 6 from Patent WO0206517.
 DEFINITION
 ACCESSION AX357080
 VERSION AX357080.1 GI:18674262
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.

REFERENCE 1
 AUTHORS Bylina, E.J., Coleman, W.J. and Youvan, D.C.
 TITLE High-throughput methods for generating and screening compounds that
 affect cell viability
 JOURNAL Patent: WO 0206517-A 6 24-JAN-2002;
 Kairos Scientific, Inc. (US)

FEATURES Location/Qualifiers
 source
 1. 68
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Nucleic acid encoding ubiquitin indolicidin fusion
 protein fragment"

ORIGIN

Alignment Scores:
 Pred. No.: 0.407 Length: 68
 Score: 74.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 90.00% Mismatches: 0
 Query Match: 86.05% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x AX357080 (1-68)

QY 2 LeuArgTrpProTrpTrpProTrpArgArg 11
 Db 28 TTGAATGGCTTGGTGGCGGTGGCGTGC 57

RESULT 6

BD128612 39 bp DNA linear PAT 18-SEP-2002
 LOCUS Method for mass production of antimicrobial peptide.
 DEFINITION
 ACCESSION BD128612
 VERSION BD128612.1 GI:23223557
 KEYWORDS JP 2002502246-A/15.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 39)
 AUTHORS Kim, S.C., Lee, J.H., Kang, M.H., Kim, J.H., Hong, S.S. and Lee, H.S.
 TITLE Method for mass production of antimicrobial peptide
 JOURNAL Patent: JP 2002502246-A 15 22-JAN-2002;
 SAMYANG GENEX CORP, KOREA ADVANCED INSTITUTE OF SCIENCE AND
 TECHNOLOGY

COMMENT OS Indolicidin gene
 FN JP 2002502246-A/15
 PD 22-JAN-2002
 PF 28-MAY-1998 JP 1999500514
 PR 28-MAY-1997 KR 1997/21312, 09-APR-1998 KR 1998/13372 PI
 SUN CHANG KIM, JAE HYUN LEE, MIN HYUNG KANG, JEONG HYUN KIM, SEUNG
 PI SUH HONG,
 PI HYUN SOO LEE
 PC C12N15/62
 CC Method for mass production of antimicrobial peptide FH Key
 Location/Qualifiers
 FT source
 1. 39
 /organism="Indolicidin gene".
 Location/Qualifiers
 1. 39
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source

ORIGIN

Alignment Scores:
 Pred. No.: 0.665 Length: 39
 Score: 70.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 88.89% Mismatches: 0
 Query Match: 81.40% Indels: 0
 DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x BD128612 (1-39)

QY 3 ArgTrpProTrpTrpProTrpArgArg 11
 Db 13 AAATGGCGGTGGTGGCGGTGGCGTGC 39

```

RESULT 7
AR404925
LOCUS AR404925 69 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 630197.
ACCESSION AR404925
VERSION AR404925.1 GI:40153719
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 69)
AUTHORS Wood, T.K., Jayaraman, A. and Earthman, J.C.
TITLE Inhibition of sulfate-reducing-bacteria-mediated degradation using
bacteria which secrete antimicrobials
JOURNAL Patent: US 630197-A 1 07-OCT-2003;
FEATURES
    source
        1..69
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.2 Length: 69
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0
US-09-444-281-36 (1-12) x AR404925 (1-69)
QY 3 ArgTTPProTTPProTTPArgArg 11
DB 28 AATGGCCTTGCTGGCTTGCGCCGC 54
RESULT 8
AR226388
LOCUS AR226388 211 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 11 from patent US 644645.
ACCESSION AR226388
VERSION AR226388.1 GI:27264888
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 211)
AUTHORS Selsted, M.E. and Osapay, K.
TITLE Crosslink-stabilized indolicidin analogs
JOURNAL Patent: US 644645-A 11 03-SEP-2002;
FEATURES
    source
        1..211
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3.79 Length: 211
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0
US-09-444-281-36 (1-12) x AR226388 (1-211)
QY 3 ArgTTPProTTPProTTPArgArg 11
DB 38 AATGGCCTTGCTGGCTTGCGCTCGT 64
RESULT 9
AR282754
LOCUS AR282754 211 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 38 from patent US 6524585.

```

```

ACCESSION AR282754
VERSION AR282754.1 GI:29719494
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 211)
AUTHORS Selsted, M.E.
TITLE Indolicidin analogs and methods of using same
JOURNAL Patent: US 6524585-A 38 25-FEB-2003;
FEATURES
    source
        1..211
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3.79 Length: 211
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0
US-09-444-281-36 (1-12) x AR282754 (1-211)
QY 3 ArgTTPProTTPProTTPArgArg 11
DB 38 AATGGCCTTGCTGGCTTGCGCTCGT 64
RESULT 10
BTINDLCD
LOCUS BTINDLCD 550 bp mRNA linear MAM 15-NOV-2001
DEFINITION Bos taurus mRNA for cathelicidin (CATHL4 gene).
ACCESSION X67340
VERSION X67340.1 GI:462
KEYWORDS cathelicidin; CATHL4 gene; indolicidin antimicrobial peptide.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1
AUTHORS Del Sal, G., Storici, P., Schneider, C., Romeo, D. and Zanetti, M.
TITLE cDNA cloning of the neutrophil bactericidal peptide indolicidin
JOURNAL Biochem. Biophys. Res. Commun. 187 (1), 467-472 (1992)
MEDLINE 92392368
PUBMED 1520337
REFERENCE 2 (bases 1 to 550)
AUTHORS Del Sal, G.
TITLE Direct Submission
JOURNAL
Biochimica, Biofisica e Chimica, delle Macromolecole, Via A.
Valerio, 38, 34127 Trieste, ITALY
FEATURES
    source
        1..550
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /cell_line="bone marrow cells"
            1..550
            /gene="CATHL4"
            13..447
            /gene="CATHL4"
            /function="indolicidin antimicrobial peptide"
            /codon_start=1
            /product="cathelicidin"
            /protein_id="CAA47755.1"
            /db_xref="GI:463"
            /db_xref="GOA:P33046"
            /db_xref="SWISS-PROT:P33046"
            /translation="MOTQPSLSLGRWSLMLLLGLVPSASQAALSYREAVLRVDQ
            INLSEANLYRLLEDPFKNEDLGRKPVSVFTYKTVCPRTIQPQPCQDFKEKG

```

sig_peptide 13..99
mat_peptide 403..444
RVKQCVGTLDPSNDQDLNCLNELQSVLPWKWPWPWRRC
/gene="CATHL4"
/product="cathelicidin"

ORIGIN

Alignment Scores:
Pred. No.: 10.2 Length: 550
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x BTINDLCD (1-550)

QY 3 ArgTTPProTTPProTTPArgArg 11
Db 415 AAATGGCCATGGTGGCTTGGCGGAGA 441

RESULT 11

ARL173324 ARL173324 6446 bp DNA linear PAT 17-DEC-2001
LOCUS
DEFINITION Sequence 5 from patent US 6303779.
ACCESSION ARL173324
VERSION ARL173324.1 GI:17912815
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 6446)
Garger,S.J., Holtz,R.Barry., McCulloch,M.J. and Turpen,T.H.
Process for isolating and purifying viruses and sugars from plant sources

JOURNAL Patent: US 6303779-A 5 16-OCT-2001;

FEATURES

source
Location/Qualifiers
1..6446
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 129 Length: 6446
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x ARL173324 (1-6446)

QY 3 ArgTTPProTTPProTTPArgArg 11
Db 6213 AAGTGGCTTGGTGGCCATGGCGGAGA 6239

RESULT 12

AX098418 AX098418 6446 bp RNA linear PAT 03-APR-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0119969.

ACCESSION AX098418

VERSION AX098418.1 GI:13537710

KEYWORDS

SOURCE

Nicotiana benthamiana

ORGANISM

Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

1
Garger,S.J., Holtz,B.R., McCulloch,M.J. and Turpen,T.H.
A process for isolating and purifying viruses, soluble proteins and peptides from plant sources

JOURNAL

Patent: WO 0119969-A 5 22-MAR-2001;

Large Scale Biology Corporation (US)

FEATURES

source

Location/Qualifiers
1..6446
/organism="Nicotiana benthamiana"
/mol_type="unassigned RNA"
/db_xref="taxon:4100"

ORIGIN

Alignment Scores:
Pred. No.: 129 Length: 6446
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x AX098418 (1-6446)

QY 3 ArgTTPProTTPProTTPArgArg 11
Db 6213 AAGTGGCTTGGTGGCCATGGCGGAGA 6239

RESULT 13

AL672300 AL672300 149991 bp DNA linear ROD 14-JUN-2002
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-446K8 on chromosome X, complete sequence.

ACCESSION AL672300

VERSION AL672300.5 GI:21436714

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 149991)

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (14-JUN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jun 17, 2002 this sequence version replaced gi:21261899.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.choi.org/bacpac/home.htm>
VECTOR: pBAC3.6.

FEATURES

source

Location/Qualifiers
1..149991
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

/chromosome="X"

/clone="RP23-446K8"

/clone_lib="RPCI-23"

ORIGIN

Alignment Scores: 3.3e+03 149991
Pred. No.: 70.00 Matches: 9
Score: 90.00% Conservative: 0
Percent Similarity: 90.00% Mismatches: 1
Best Local Similarity: 90.00% Indels: 0
Query Match: 81.40% Gaps: 0
DB: 10

US-09-444-281-36 (1-12) x AL672300 (1-149991)

QY 1 IleLeuArgTrrProTrrTrrProTrrArg 10
DB 12272 ATATGGAGTGGCCATGGTGGCCCTGGCCG 12301

RESULT 14
AC113900 195764 bp DNA linear HTG 15-NOV-2002
LOCUS
DEFINITION *** 2 unordered pieces.
AC113900
AC113900.5 GI:25007914
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
RATTUS NORVEGICUS (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Muridae; Murinae;
Rattus.
1 (bases 1 to 195764)
Muzny,D.Marie., Metaker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooke,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buahy,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,M., Hamil,C., Hamilton,N., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,M.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowitz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuwa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Manghwar,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Narkervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smailes,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Streimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villanasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

site:
end_sequence:BZ164864"
109081..110233
/note="wgs_contig"

misc_feature

ORIGIN

Alignment Scores: 4.34e+03 195764
Pred. No.: 70.00 Matches: 8
Score: 100.00% Conservative: 1
Percent Similarity: 88.89% Mismatches: 0
Best Local Similarity: 81.40% Indels: 0
Query Match: 2 Gaps: 0
DB:

US-09-444-281-36 (1-12) x AC113900 (1-195764)

Qy 3 ArgTTPProTTPTrpProTTPArgArg 11

Db 187911 AGGTGGCCATGGTGGCCATGGAGCG 187937

RESULT 15

AC095460 255121 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-7J8, WORKING DRAFT SEQUENCE.
DEFINITION AC095460
AC095460.6 GI:30467723
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 255121)
Munzy,D,Warrie, Metzker,M, Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,I.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newcom,N., Nguyen,N., Norris,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-I.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,K., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,

Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tinney,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Yoon,L., Yoon,V.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 255121)

Worley,K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 255121)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24941100.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GCJK

Center clone name: CH230-7J8

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 243844 bases at least Q40

Consensus quality: 245664 bases at least Q30

Consensus quality: 246956 bases at least Q20

Estimated insert size: 260749; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 255121: contig of 255121 bp in length.

Location/Qualifiers

1..255121

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-7J8"

1..1163

/note="wgs_contig"

FEATURES

source

misc_feature

```
misc_feature 252824..255121
/note="wgs_contig"

ORIGIN
Alignment Scores:
Pred. No.:      5.7e+03      Length:      255121
Score:          70.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 88.89%      Mismatches:  0
Query Match:      81.40%      Indels:      0
DB:              2          Gaps:      0

US-09-444-281-36 (1-12) x AC095460 (1-255121)

QY      3 ArgTTPProTTPProTTPArgArg 11
      |||||
Db      42133 AGTGGCCATGGTGGCCATGAAGCGG 42159

Search completed: May 11, 2004, 16:01:51
Job time : 2243.92 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 12:39:07 ; Search time 237.12 Seconds

(without alignments)
214.990 Million cell updates/sec

Title: US-09-444-281-36

Perfect score: 86

Sequence: 1 ILRWPMWRRK 12

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cpn2.1/USPTO.spool/US09444281/runat 07052004 171138 2561/app query.fasta_1.398
-DB=N_Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281 -CGN 1 1 586 -runat 07052004 171138 2561 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 29Jan04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	108	3	AAA27296
2	86	100.0	114	3	AAA27298
3	86	100.0	151	3	AAA27294
4	75	87.2	88	2	AAV60908
5	75	87.2	114	3	AAA27291
6	74	86.0	68	6	AAI72481
7	70	81.4	39	2	AAV83788
8	70	81.4	47	3	AAZ29389

c	9	70	81.4	47	3	AAZ29390
c	10	70	81.4	47	6	ABL60414
c	11	70	81.4	47	6	ABL60415
c	12	70	81.4	53	3	AAZ29364
c	13	70	81.4	53	6	ABL60445
c	14	70	81.4	69	3	AAZ40246
c	15	70	81.4	211	3	AAZ49764
c	16	70	81.4	211	3	AAZ45123
c	17	70	81.4	550	6	AAZ45350
c	18	70	81.4	6446	2	AAZ20646
c	19	70	81.4	6446	2	AAZ20646
c	20	67	77.9	207	3	AAA28519
c	21	67	77.9	1040	6	ABL58680
c	22	67	77.9	1040	6	ABA96630
c	23	65	75.6	39	3	AAA28510
c	24	65	75.6	39	9	ADC73335
c	25	65	75.6	52	3	AAA28514
c	26	65	75.6	52	3	AAA28512
c	27	65	75.6	53	9	ADC73337
c	28	65	75.6	54	9	ADC73339
c	29	64.5	75.0	1539	7	ADA70241
c	30	63	73.3	1437	4	AAZ54129
c	31	63	73.3	1848	4	ABL21537
c	32	63	73.3	1908	4	AAZ56079
c	33	63	73.3	1908	7	ACA51647
c	34	63	73.3	4433	4	ABL21536
c	35	63	73.3	6957	9	ADB52723
c	36	62	72.1	48	3	AAA27290
c	37	62	72.1	49	3	AAA27287
c	38	61	70.9	758	2	AAQ99309
c	39	61	70.9	837	5	AAZ79653
c	40	61	70.9	1065	5	AAZ75627
c	41	61	70.9	1291	4	ABL15825
c	42	61	70.9	1676	3	AAZ46155
c	43	61	70.9	2034	9	ADC10047
c	44	61	70.9	2110	3	AAZ46160
c	45	61	70.9	2213	4	ABL22542

ALIGNMENTS

RESULT 1

AAA27296

ID AAA27296 standard; DNA; 108 BP.

XX

AC AAA27296;

XX

DT 20-SEP-2000 (first entry)

XX

DE Oligonucleotide used for synthesis of MBI 2X1187 poly cassette.

XX

DE Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11;

XX

KW indolicidin; bovine; ss.

XX

OS Synthetic.

XX

FN W0200031279-A2.

XX

PD 02-JUN-2000.

XX

PF 19-NOV-1999; 99WO-CA001107.

XX

PR 20-NOV-1998; 98US-0109218P.

XX

PA (MICR-) MICROLOGIX BIOTECH INC.

XX

PI Burian J, Bartfeld D;

XX

DR WPI; 2000-400086/34.

XX

PT Multi-domain fusion protein expression cassette used for high yield

PT stable production of foreign peptide gene products.

XX

PS Example 5; Page 39; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in recombinant host cells. The method involves construction of a multi-domain fusion protein expression cassette comprising a promoter and a nucleic acid molecule expressed as an insoluble protein. The inclusion of anionic peptide sequences in the linker sequences neutralises the positive charge of the cationic peptide so that the charge of the fusion protein is controlled. This cassette allows high yield, stable production of the cationic peptide. Cationic peptides such as bovine indolicidin may be used as antimicrobial agents. The present sequence is an oligonucleotide that was used to synthesise a MBI-11B7 fragment. This fragment was used in the expression cassette. MBI-11B7 is a cationic peptide derived from modifications of indolicidin

XX SQ Sequence 108 BP; 18 A; 33 C; 31 G; 26 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.0254	Length:	108
Score:	86.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-444-281-36 (1-12) x AAA27296 (1-108)

QY 1 lleLeuArgTrpProTrpArgLys 12

Db 41 ATTCTGGTTGGCGGTGGTGGCGTGGCGTGCACAA 76

RESULT 2

AAA27298

ID AAA27298 standard; DNA; 114 BP.

AC AAA27298;

XX 20-SEP-2000 (first entry)

XX Oligonucleotide used for synthesis of MBI 11B7 first cassette.

XX Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11; indolicidin; bovine; ss.

XX Synthetic.

XX WO200031279-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MTCR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MTCR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX Multi-domain fusion protein expression cassette used for high yield stable production of foreign peptide gene products.

XX Example 5; Page 40; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in recombinant host cells. The method involves construction of a multi-domain fusion protein expression cassette comprising a promoter and a nucleic acid molecule expressed as an insoluble protein. The inclusion of anionic peptide sequences in the linker sequences neutralises the positive charge of the cationic peptide so that the charge of the fusion protein is controlled. This cassette allows high yield, stable production of the cationic peptide. Cationic peptides such as bovine indolicidin may

CC be used as antimicrobial agents. The present sequence is an oligonucleotide that was used to synthesise a MBI-11B7 fragment. This fragment was used in the expression cassette. MBI-11B7 is a cationic peptide derived from modifications of indolicidin

XX SQ Sequence 114 BP; 20 A; 34 C; 32 G; 28 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.0269	Length:	114
Score:	86.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-444-281-36 (1-12) x AAA27298 (1-114)

QY 1 lleLeuArgTrpProTrpArgLys 12

Db 47 ATTCTGGTTGGCGGTGGTGGCGTGGCGTGCACAA 82

RESULT 3

AAA27294

ID AAA27294 standard; DNA; 151 BP.

AC AAA27294;

XX 20-SEP-2000 (first entry)

XX Oligonucleotide used for synthesis of MBI 2X11B7 last cassette.

XX Oligonucleotide; cellulose binding domain; CBD; cationic peptide; MBI-11; indolicidin; bovine; ss.

XX Synthetic.

XX WO200031279-A2.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-CA001107.

XX 20-NOV-1998; 98US-0109218P.

XX (MTCR-) MICROLOGIX BIOTECH INC.

XX Burian J, Bartfeld D;

XX WPI; 2000-400086/34.

XX Multi-domain fusion protein expression cassette used for high yield stable production of foreign peptide gene products.

XX Example 5; Page 38; 73pp; English.

XX A novel method allows the efficient production of cationic peptides in recombinant host cells. The method involves construction of a multi-domain fusion protein expression cassette comprising a promoter and a nucleic acid molecule expressed as an insoluble protein. The inclusion of anionic peptide sequences in the linker sequences neutralises the positive charge of the cationic peptide so that the charge of the fusion protein is controlled. This cassette allows high yield, stable production of the cationic peptide. Cationic peptides such as bovine indolicidin may be used as antimicrobial agents. The present sequence is an oligonucleotide that was used to synthesise a MBI-11B7 fragment. This fragment was used in the expression cassette. MBI-11B7 is a cationic peptide derived from modifications of indolicidin

XX SQ Sequence 151 BP; 22 A; 44 C; 49 G; 36 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.0361	Length:	151
Score:	86.00	Matches:	12

XX DE cDNA derived from C-terminus of ubiquitin.

XX KW Recursive ensemble mutagenesis; REM, cell viability; optical signal;

XX KW high-throughput screening; antimicrobial compound; antibiotic; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT CDS 1..60

XX FT /*tag= a

XX FT /partial

XX FT /transl_except= pos:28..30, aa:Trp

XX FT /note= "No start codon given"

XX PN WO200206517-A2.

XX PD 24-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US023004.

XX PR 19-JUL-2000; 2000US-0219179P.

XX PA (KAIR-) KAIROS SCI INC.

XX PI Bylina EJ, Coleman WJ, Youvan DC;

XX PI WPI; 2002-179801/23.

XX DR P-PSDB; AAB47907.

XX Screening compounds affecting cell viability e.g. for identifying

PT antimicrobial compounds, comprises determining if induced transformed

PT cell colonies have a desired signal when contacted with a viability

PT indicator.

XX PS Disclosure; Fig 6; 56pp; English.

XX This sequence is derived from the C-terminus of ubiquitin and represents

CC the cloning region. SacII-BglII cassettes can be used for inserting

CC antimicrobial peptide sequences into this region. This sequence may be

CC used to form the recursive ensemble mutagenesis (REM) cassette of the

CC invention. The cassette may be used in the method of the invention for

CC determining whether a compound affects cell viability by: (a) exposing

CC colonies of cells (CC) on a support surface to inducing conditions, where

CC the cells have been transformed with an expression library encoding

CC candidate compounds; (b) contacting CC with a viability indicator that

CC produces an optical signal indicative of cell viability; and (c)

CC determining if a colony has a desired optical signal. The method is

CC useful for determining whether a compound affects cell viability. It is

CC useful for high-throughput screening to identify antimicrobial compounds

CC and in drug discovery. The antimicrobial compounds are useful in the

CC pharmaceutical industry, and provide an additional new class of

CC antibiotic compounds to fight infectious diseases. The method is useful

CC for assaying the authentic peptide sequences contained in an expression

CC library for antimicrobial activity, for distinguishing dead cells

CC (expressing active sequences) from living cells (expressing inactive or

CC less active sequences), to identify novel antimicrobial peptide

CC sequences, including highly potent molecules, resulting in a large number

CC of new antimicrobial lead compounds that are active against a broad range

CC of bacteria or other microorganisms, and for screening all types of

CC antibiotic compounds, including libraries of low molecular weight

CC molecules produced by metabolic engineering and artificial synthesized

CC libraries in solid-phase arrays

XX SQ Sequence 68 BP; 11 A; 18 C; 19 G; 20 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.398 Length: 68

Score: 74.00 Matches: 9

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 90.00% Mismatches: 0

Query Match: 86.05% Indels: 0

DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x AAI72481 (1-68)

QY 2 LeuArgTrpProTrpTrpProTrpArgArg 11

Db 28 TTGAATGGCTTGGTGGCTTGGCGTCGC 57

RESULT 7

AAV83788

ID AAV83788 standard; DNA; 39 BP.

XX

AC AAV83788;

XX

DT 19-MAR-1999 (first entry)

XX

DE Antimicrobial peptide Indolicidin encoding DNA.

XX

KW Antimicrobial; fusion; acidic peptide; recombinant; microorganism;

KW guamerin; basic peptide; Indolicidin; ss.

XX

OS Synthetic.

OS Bos sp.

XX

FH Key Location/Qualifiers

FT CDS 1..39

FT /*tag= a

FT /note= "the start and stop codons are not indicated"

XX

PN WO9854336-A1.

XX

PD 03-DEC-1998.

XX

PF 28-MAY-1998; 98WO-KR000132.

XX

PR 28-MAY-1997; 97KR-00021312.

PR 09-APR-1998; 98KR-00013372.

XX

XX (SAMY-) SAMYANG GENEX CORP.

PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.

XX

PI Kim S, Lee JH, Kang MH, Kim JH, Hong S, Lee H;

XX

WPI; 1999-059844/05.

DR P-PSDB; AAM87609.

XX

PT New method for mass production of antimicrobial peptides - by

PT constructing fusion genes comprising acidic and antimicrobial peptide

PT genes and transforming host with vector containing these.

XX

Example 6; Page 18; 52pp; English.

XX

CC The invention relates to mass production of antimicrobial peptides. The

CC method comprises constructing a fusion gene containing a first gene

CC encoding a negatively charged acidic peptide having at least two cysteine

CC residues, and a second gene encoding a positively charged basic

CC antimicrobial peptide. A host microorganism is transformed with a vector

CC containing the fusion gene and then cultured. The expressed antimicrobial

CC peptide is then recovered. The method is used to mass produce

CC antimicrobial peptides in recombinant microorganisms. The inhibitory

CC effect of the expressed antimicrobial peptide upon the growth of the host

CC microorganism is considerably reduced by fusing it to the acidic peptide.

CC Therefore, the use of the fusion gene provides an economic, recombinant

CC alternative of mass producing antimicrobial peptides, which overcomes the

CC disadvantages of low-productivity and poor economy, previously

CC encountered by recombinant and chemical methods. The present sequence

CC represents the DNA encoding an antimicrobial peptide Indolicidin. This

CC can be used along with the acidic peptide Guamerin gene in the

CC construction of the fusion gene

XX SQ Sequence 39 BP; 4 A; 10 C; 16 G; 9 T; 0 U; 0 Other;

Alignment Scores: 0.654 Length: 39

Pred. No.: 0.654

DE DNA fragment of the invention #15.
XX
KW Gene expression; peptide antibiotic; purF gene; ds.
XX
OS Unidentified.
XX
PN KR2001098973-A.
XX
PD 08-NOV-2001.
XX
PF 08-JUN-2001; 2001KR-00031889.
XX
PR 08-JUN-2001; 2001KR-00031889.
XX
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;
XX
XX WPI; 2002-301977/34.
XX
PT Gene expression system useful for mass-production of peptide antibiotics
PT and vectors derived from microorganisms.
XX
PS Disclosure; Page 7; 56pp; Korean.
XX
CC The invention relates to a genetic component which mass-produces peptide
CC antibiotics effectively from microorganisms. The genetic component
CC consists of a first gene sequence which codes for the whole or partial
CC purF gene or its derivative, and a second gene sequence which codes
CC peptide antibiotics. The mass-production method of peptide antibiotics
CC comprises the steps of; constructing an expression vector including the
CC genetic component, transforming a bacterial host cell with the vector,
CC culturing the transformed cell to express the genetic component, and
CC recovering the peptide antibiotics. The expression vector is selected
CC from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a
CC high copy number of origin, strong transcription promoter and structural
CC gene. The sequences given in records ABL60400-ABL60464 represent DNA
CC sequences of the invention
XX
SQ Sequence 47 BP; 6 A; 11 C; 19 G; 11 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.795 Length: 47
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.99% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x ABL60414 (1-47)

QY 3 ArgTtpProTtpTtpProTtpArgArg 11
Db 17 AAATGCCGCTGGTGGCGGTGGTGGT 43

RESULT 11
ABL60415/c
ID ABL60415 standard; DNA; 47 BP.
XX
AC ABL60415;
XX
XX 28-MAR-2003 (first entry)
XX
DE DNA fragment of the invention #16.
XX
KW Gene expression; peptide antibiotic; purF gene; ds.
XX
OS Unidentified.
XX
PN KR2001098973-A.
XX
PD 08-NOV-2001.
XX

PF 08-JUN-2001; 2001KR-00031889.
XX
PR 08-JUN-2001; 2001KR-00031889.
XX
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;
XX
XX WPI; 2002-301977/34.
XX
PT Gene expression system useful for mass-production of peptide antibiotics
PT and vectors derived from microorganisms.
XX
PS Disclosure; Page 7; 56pp; Korean.
XX
CC The invention relates to a genetic component which mass-produces peptide
CC antibiotics effectively from microorganisms. The genetic component
CC consists of a first gene sequence which codes for the whole or partial
CC purF gene or its derivative, and a second gene sequence which codes
CC peptide antibiotics. The mass-production method of peptide antibiotics
CC comprises the steps of; constructing an expression vector including the
CC genetic component, transforming a bacterial host cell with the vector,
CC culturing the transformed cell to express the genetic component, and
CC recovering the peptide antibiotics. The expression vector is selected
CC from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has a
CC high copy number of origin, strong transcription promoter and structural
CC gene. The sequences given in records ABL60400-ABL60464 represent DNA
CC sequences of the invention
XX
SQ Sequence 47 BP; 12 A; 18 C; 10 G; 7 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.795 Length: 47
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.99% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 6 Gaps: 0

US-09-444-281-36 (1-12) x ABL60415 (1-47)

QY 3 ArgTtpProTtpTtpProTtpArgArg 11
Db 35 AAATGCCGCTGGTGGCGGTGGTGGT 9

RESULT 12
AAZ29364
ID AAZ29364 standard; DNA; 53 BP.
XX
AC AAZ29364;
XX
XX 29-FEB-2000 (first entry)
XX
DE Antimicrobial peptide, Indolicidin encoding DNA.
XX
KW purF gene; glutamine pyrophosphoribosyl pyrophosphate amidotransferase;
KW purF derivative; fusion partner; antimicrobial peptide; Indolicidin;
KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct; cow;
KW neutralise; toxicity; pharmaceutical industry; food industry; ds.
XX
OS Bos taurus.
XX
XX Key Location/Qualifiers
FH CDS 5..46
FT /*tag= a
FT /product= "Indolicidin peptide"
FT /note= "Antimicrobial peptide used in DNA construct"
XX
XX WO9964611-A1.
XX
XX 16-DEC-1999.
XX
XX 08-JUN-1999; 99WO-KR000282.
XX
PF

XX 09-JUN-1998; 98XR-00022117.
 PR 14-MAY-1999; 99XR-00017920.
 XX (SAMY-) SAMYANG GENEX CORP.
 PA Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;
 PI WPI; 2000-097542/08.
 DR P-PSDB; AAY44324.
 XX New DNA constructs useful for mass production of antimicrobial peptides
 PT in microorganism hosts.
 CC Claim 1; Fig 1; 67pp; English.
 XX The present DNA sequence encodes an antimicrobial peptide, Indolicidin
 CC derived from cow, Bos taurus. It is used along with a derivative of purF
 CC gene sequence that functions as a fusion partner. A DNA construct that
 CC comprises, this antimicrobial peptide encoding sequence and the entire,
 CC partial or derivative of purF gene, is used for mass production of the
 CC antimicrobial peptide in microorganisms without killing the host cells.
 CC Use of the purF gene derivative sequence, neutralises the toxicity of the
 CC antimicrobial peptides against the host microorganism. The antimicrobial
 CC peptides are useful commercially in the pharmaceutical and food
 CC industries
 XX SQ Sequence 53 BP; 8 A; 12 C; 20 G; 13 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.902 Length: 53
 Score: 70.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 88.89% Mismatches: 0
 Query Match: 81.40% Indels: 0
 DB: 3 Gaps: 0
 US-09-444-281-36 (1-12) x AAZ29364 (1-53)
 Qy 3 ArgTrrProTrrPrrProTrrArg 11
 Db 17 AAATGGCGTGGTGGCGTGGCGTGGT 43
 RESULT 13
 ABL60445
 ID ABL60445 standard; DNA; 53 BP.
 XX AC ABL60445;
 XX 28-MAR-2003 (first entry)
 DT DNA fragment of the invention #44.
 DE Gene expression; peptide antibiotic; purF gene; ds.
 XX Unidentified.
 XX Key Location/Qualifiers
 FT CDS 5..60
 FT /*tag= a
 FT /partial
 FT /note= "no start codon present"
 XX KB2001098973-A.
 XX 08-NOV-2001.
 XX 08-JUN-2001; 2001KR-00031889.
 XX 08-JUN-2001; 2001KR-00031889.
 XX (SAMY-) SAMYANG GENEX CORP.
 XX

PI Hong SS, Kang MH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;
 XX WPI; 2002-301977/34.
 DR P-PSDB; ABB81940.
 XX Gene expression system useful for mass-production of peptide antibiotics
 PT and vectors derived from microorganisms.
 CC Disclosure; Page 17; 56pp; Korean.
 XX The invention relates to a genetic component which mass-produces peptide
 CC antibiotics effectively from microorganisms. The genetic component
 CC consists of a first gene sequence which codes for the whole or partial
 CC purF gene or its derivative, and a second gene sequence which codes
 CC peptide antibiotics. The mass-production method of peptide antibiotics
 CC comprises the steps of; constructing an expression vector including the
 CC genetic component, transforming a bacterial host cell with the vector,
 CC culturing the transformed cell to express the genetic component, and
 CC recovering the peptide antibiotics. The expression vector is selected
 CC from the group consisting of pGMX2, pGMX3, pGMX4 and pGMX5, and it has a
 CC high copy number of origin, strong transcription promoter and structural
 CC gene. The sequences given in records ABL60400-ABL60464 represent DNA
 CC sequences of the invention
 XX SQ Sequence 53 BP; 8 A; 12 C; 20 G; 13 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.902 Length: 53
 Score: 70.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 88.89% Mismatches: 0
 Query Match: 81.40% Indels: 0
 DB: 6 Gaps: 0
 US-09-444-281-36 (1-12) x ABL60445 (1-53)
 Qy 3 ArgTrrProTrrPrrProTrrArg 11
 Db 17 AAATGGCGTGGTGGCGTGGCGTGGT 43
 RESULT 14
 AAZ40246
 ID AAZ40246 standard; DNA; 69 BP.
 XX AC AAZ40246;
 XX 23-FEB-2000 (first entry)
 DT Oligonucleotide for cloning indolicidin peptide coding sequence.
 DE Indolicidin; bactericin; sulphate-reducing bacteria; growth inhibitor;
 KW corrosion; degradation; metal; concrete; cement; dental implant; biofilm;
 KW ss.
 XX Synthetic.
 OS Bacillus sp.
 XX WO9956553-A1.
 XX 11-NOV-1999.
 XX 03-MAY-1999; 99WO-US009675.
 XX 06-MAY-1998; 98US-00074037.
 XX 31-MAR-1999; 99US-00282277.
 XX (REGC) UNIV CALIFORNIA.
 XX Wood TK, Jayaraman A, Earthman JC;
 XX WPI; 2000-052882/04.
 XX Inhibiting growth of sulfate-reducing bacteria using other bacteria,
 PT

FT particularly for protection of metals and concrete.

PS Example 4; Fig 1; 84pp; English.

XX This sequence represents an oligonucleotide for cloning the non-amidated
 CC indolicidin peptide coding sequence. The invention relates to a method
 CC for inhibiting growth of sulphate-reducing bacteria (A) on a material (B)
 CC sensitive to corrosion or degradation, by applying to (B) a bacterium (C)
 CC that secretes a compound (I) able to inhibit growth of (A). The method is
 CC used to protect metal, concrete or cement against corrosion and
 CC degradation, but (B) can also be used to protect dental implants, (B) is
 CC present in an open or closed system (e.g. water cooling tower, liquid
 CC storage container, fuel tank, sewer or drainage system etc.) or part of a
 CC bridge or other structure. The method is more effective and less
 CC expensive than known methods for inhibiting (A), and reduces the amount
 CC of toxic chemicals released. Conventional biofilms of aerobic organisms
 CC tend to encourage growth of (A), and addition of (C) to the biofilm
 CC prevents this. A single application of (C) lasts for a long time, and (I)
 CC are produced exactly where they are required and inhibit (A) without
 CC significant impact on other organisms (this effect includes reducing
 CC resistance of (A) to conventional biocides, which may then be used in
 CC reduced amounts). If local damage to the biofilm occurs, the underlying
 CC material is still protected by diffusion of (I) from neighbouring areas
 XX
 SQ Sequence 69 BP; 14 A; 18 C; 20 G; 17 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.19 Length: 69
 Score: 70.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 88.89% Mismatches: 0
 Query Match: 81.40% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-36 (1-12) x AAZ40246 (1-69)

QY 3 ArgTrpProTrpTrpProTrpArgArg 11

DB 28 AAATGGCGTGGTGGCGTGGCGCGC 54

RESULT 15

AAZ49764

ID AAZ49764 standard; DNA; 211 BP.

XX AC AAZ49764;

XX DT 18-APR-2000 (first entry)

XX DE Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met) 3 DNA.

XX KW Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13;
 KW stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
 KW protozoacide; virucide; anti-HIV; human immunodeficiency virus-1; HIV-1;
 KW gram positive bacteria; gram negative; Staphylococcus aureus;
 KW Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
 KW Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;
 KW hexapeptide spacer; ds.

XX OS Synthetic.

XX OS Bos sp.

XX FH Key Location/Qualifiers

FT primer_bind 1..21

FT CDS 8..199

FT /*tag= b

FT /*tag= a

FT /product= "Poly-(Indol(1-13)-Met-Ala-Arg-Ile-Ala-Met)3"

FT /note= "encodes three copies of Indol 1-13, each

FT separated by Met-Ala-Arg-Ile-Ala-Met spacer sequence"

FT 68..71

FT misc_feature

FT /*tag= d

FT /note= "corresponds to overlap in oligonucleotides used

FT for ligation"

FT misc_feature 148..151

FT /*tag= e

FT /note= "corresponds to overlap in oligonucleotides used

FT for ligation"

FT primer_bind complement(191..211)

FT /*tag= c

XX WO9965510-A1.

XX 23-DEC-1999.

XX 20-MAY-1999; 99WO-US011165.

XX 18-JUN-1998; 98US-0009631.

XX (REGC) UNIV CALIFORNIA.

XX Seisted ME, Osapay K;

XX WPI; 2000-147133/13.

XX P-PSDB; AAZ44668.

XX Crosslinked indolicidin analogs with antimicrobial activity against

XX bacteria, yeast, fungi, protozoa and viruses.

XX Example 1C; Fig 1; 53pp; English.

XX The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
 CC which is a naturally occurring peptide isolated from bovine neutrophils
 CC and has antimicrobial activity. The crosslinked indolicidin (X-
 CC indolicidin) analogs are stable and have antimicrobial activity against
 CC gram positive and negative bacteria (e.g. Staphylococcus aureus,
 CC Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
 CC Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
 CC species and Acanthamoeba species), and viruses (e.g. HIV-1). They can be
 CC used for reducing or inhibiting the growth or survival of microorganisms
 CC in an environment e.g. a food or food product, a solution, an inanimate
 CC object comprising a surface, or a mammal. The present sequence is a DNA
 CC encoding a protein comprising three copies of Indol 1-13 each separated
 CC by a hexapeptide spacer sequence. The sequence was used to produce a
 CC recombinant construct for the expression of Indol-homoserine (Hee)
 CC analog. The ability of Indol-Hse analog to maintain antimicrobial
 CC activity provides a means to produce X-indolicidin analog precursors in
 CC sufficient quantities

XX SQ Sequence 211 BP; 36 A; 50 C; 74 G; 51 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.84 Length: 211
 Score: 70.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 88.89% Mismatches: 0
 Query Match: 81.40% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-36 (1-12) x AAZ49764 (1-211)

QY 3 ArgTrpProTrpTrpProTrpArgArg 11

DB 38 AAATGGCGTGGTGGCGTGGCGTGGT 64

Search completed: May 11, 2004, 14:43:52

Job time : 238.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:02:47 ; Search time 46.08 Seconds
(without alignments)
144,518 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRNPWPWRK 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09444281/runat 07052004 171139 2589/app query.fasta_1.398
-DB=Issued Patents NA -QWMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281 @cgn_1_176 @runat 07052004 171139 2589 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	87.2	88	3	US-09-042-071-49
2	70	81.4	39	3	Sequence 49, Appl
3	70	81.4	69	4	Sequence 29, Appl
4	70	81.4	211	4	Sequence 1, Appl
5	70	81.4	211	4	Sequence 11, Appl
6	70	81.4	6446	3	Sequence 38, Appl
7	70	81.4	6446	3	Sequence 5, Appl
8	70	81.4	6446	3	Sequence 5, Appl
9	63	73.3	1378	4	Sequence 52969, Ap
10	63	73.3	1378	4	Sequence 9298, Ap
11	63	73.3	1521	4	Sequence 9281, Ap
12	61	70.9	758	3	Sequence 6, Appl

13	59	68.6	382	4	US-09-118-554-59	Sequence 59, Appl
14	59	68.6	382	4	US-09-118-627-59	Sequence 59, Appl
15	59	68.6	382	4	US-09-602-877A-59	Sequence 59, Appl
16	59	68.6	383	3	US-09-020-956-82	Sequence 82, Appl
17	59	68.6	383	3	US-09-030-607-82	Sequence 82, Appl
18	59	68.6	383	4	US-09-439-313-82	Sequence 82, Appl
19	59	68.6	383	4	US-09-352-616A-82	Sequence 82, Appl
20	59	68.6	383	4	US-09-232-149A-82	Sequence 82, Appl
21	59	68.6	383	4	US-09-159-812-82	Sequence 82, Appl
22	59	68.6	383	4	US-09-636-215-82	Sequence 82, Appl
23	59	68.6	383	4	US-09-685-166A-82	Sequence 82, Appl
24	59	68.6	383	4	US-09-115-453-82	Sequence 82, Appl
25	59	68.6	383	4	US-09-688-489-82	Sequence 82, Appl
26	59	68.6	384	3	US-09-030-607-183	Sequence 183, App
27	59	68.6	384	4	US-09-439-313-183	Sequence 183, App
28	59	68.6	384	4	US-09-352-616A-183	Sequence 183, App
29	59	68.6	384	4	US-09-232-149A-183	Sequence 183, App
30	59	68.6	384	4	US-09-159-812-183	Sequence 183, App
31	59	68.6	384	4	US-09-636-215-183	Sequence 183, App
32	59	68.6	384	4	US-09-685-166A-183	Sequence 183, App
33	59	68.6	384	4	US-09-115-453-183	Sequence 183, App
34	59	68.6	384	4	US-09-688-489-183	Sequence 183, App
35	59	68.6	440	4	US-09-118-554-45	Sequence 45, Appl
36	59	68.6	440	4	US-09-118-627-45	Sequence 45, Appl
37	59	68.6	440	4	US-09-602-877A-45	Sequence 45, Appl
38	59	68.6	499	3	US-09-020-956-73	Sequence 73, Appl
39	59	68.6	499	3	US-09-439-313-73	Sequence 73, Appl
40	59	68.6	499	4	US-09-352-616A-73	Sequence 73, Appl
41	59	68.6	499	4	US-09-232-149A-73	Sequence 73, Appl
42	59	68.6	499	4	US-09-159-812-73	Sequence 73, Appl
43	59	68.6	499	4	US-09-636-215-73	Sequence 73, Appl
44	59	68.6	499	4	US-09-685-166A-73	Sequence 73, Appl
45	59	68.6	499	4	US-09-115-453-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1

US-09-042-071-49
; Sequence 49, Application US/09042071
; Patent No. 6294372
; GENERAL INFORMATION:
; APPLICANT: Kay, William W.
; TITLE OF INVENTION: REPLICATION GENES AND GENE PRODUCTS FROM
; TITLE OF INVENTION: SMALL CRYPTIC PLASMIDS AND METHODS FOR CONSTRUCTING
; TITLE OF INVENTION: CONTROLLED-REPLICATION PLASMID VECTORS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,071
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McWaters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 660081.407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 88 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-042-071-49

Alignment Scores:
Pred. No.: 0.129 Length: 88
Score: 75.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.00% Mismatches: 0
Query Match: 87.21% Indels: 0
DB: 3 Gaps: 0

US-09-444-281-36 (1-12) x US-09-042-071-49 (1-88)

Qy 3 ArgTrpProTrpProTrpArgArgLys 12
   ::::::::::::::::::::::::::::
Db 34 AAATGGCGCGTGGTGGCGTGGCGTGTAA 63

RESULT 2
US-09-230-180-29
; Sequence 29, Application US/09230180
; Patent No. 6183992
; GENERAL INFORMATION:
; APPLICANT: Kim, Sun-Chang
; APPLICANT: Lee, Jae Hyun
; APPLICANT: Kang, Min Hyung
; APPLICANT: Kim, Jeong Hyun
; APPLICANT: Hong, Seung-Suh
; APPLICANT: Lee, Hyun-Soo
; APPLICANT: Samyang Genex Corporation
; APPLICANT: Korea Advanced Institute of Science and Technology
; TITLE OF INVENTION: METHOD FOR MASS PRODUCTION OF
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDE
; FILE REFERENCE: 6181/0F135
; CURRENT APPLICATION NUMBER: US/09/230,180
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: PCT/KR98/00132
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: KR 13372/1998
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: KR 21312/1997
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence deduced from Indolicidin peptide
; OTHER INFORMATION: sequence based on codon usage of E. coli
US-09-230-180-29

Alignment Scores:
Pred. No.: 0.209 Length: 39
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 3 Gaps: 0

US-09-444-281-36 (1-12) x US-09-230-180-29 (1-39)

Qy 3 ArgTrpProTrpProTrpArgArg 11
   ::::::::::::::::::::::::::::
Db 13 AAATGGCGCGTGGTGGCGTGGCGTGT 39

RESULT 3
US-09-282-277-1
; Sequence 1, Application US/09282277
; Patent No. 6630197

```

```

; GENERAL INFORMATION:
; APPLICANT: Wood, Thomas K.
; APPLICANT: Jayaraman, Arul
; APPLICANT: Earthman, James C.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Inhibition of Sulfate-Reducing-Bacteria-Mediated
; TITLE OF INVENTION: Degradation Using Bacteria Which Secrete Antimicrobials
; FILE REFERENCE: 02307E-0859100S
; CURRENT APPLICATION NUMBER: US/09/282,277
; CURRENT FILING DATE: 1999-03-31
; EARLIER APPLICATION NUMBER: US 09/074,037
; EARLIER FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
; OTHER INFORMATION: used for cloning indolicidin
US-09-282-277-1

Alignment Scores:
Pred. No.: 0.376 Length: 69
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x US-09-282-277-1 (1-69)

Qy 3 ArgTrpProTrpProTrpArgArg 11
   ::::::::::::::::::::::::::::
Db 28 AAATGGCGTGGTGGCGTGGCGCGCG 54

RESULT 4
US-09-099-631A-11
; Sequence 11, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
; APPLICANT: Seilsted, Michael E.
; APPLICANT: Osapay, Klara
; TITLE OF INVENTION: Crosslink-Stabilized Indolicidin Analogs
; FILE REFERENCE: P-UC 3050
; CURRENT APPLICATION NUMBER: US/09/099,631A
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)..(196)
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-099-631A-11

Alignment Scores:
Pred. No.: 1.19 Length: 211
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x US-09-099-631A-11 (1-211)

Qy 3 ArgTrpProTrpProTrpArgArg 11
   ::::::::::::::::::::::::::::

```

```
Db      38 AAATGGCGTGTGGCGGTGGCGTGTGT 64
RESULT 5
US-09-416-481A-38
; Sequence 38, Application US/09416481A
; Patent No. 6524585
; GENERAL INFORMATION:
; APPLICANT: Selsel, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3794
; CURRENT APPLICATION NUMBER: US/09/416,481A
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 09/076,227
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 38
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)..(196)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-416-481A-38
Alignment Scores:
Pred. No.:      1.19      Length:      211
Score:          70.00     Matches:      8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match:    81.40%   Indels:      0
DB:             4        Gaps:        0

US-09-444-281-36 (1-12) x US-09-416-481A-38 (1-211)
QY      3 ArgTTPProTTPProTTPArgArg 11
Db      38 AAATGGCGTGTGGCGGTGGCGTGTGT 64
RESULT 6
US-09-259-741-5
; Sequence 5, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751

; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8100
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLERCULE TYPE: Genomic RNA
US-09-259-741-5
Alignment Scores:
Pred. No.:      39.8      Length:      6446
Score:          70.00     Matches:      8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match:    81.40%   Indels:      0
DB:             3        Gaps:        0

US-09-444-281-36 (1-12) x US-09-259-741-5 (1-6446)
QY      3 ArgTTPProTTPProTTPArgArg 11
Db      6213 AAGUGGCCUUGUGGCAUGGCGCGCA 6239
RESULT 7
US-09-037-751-5
; Sequence 5, Application US/09037751
; Patent No. 6037456
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,751
; FILING DATE: 10-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
```

Alignment Scores:

Score: 63.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.26% Indels: 0
 DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x US-09-252-991A-9298 (1-1362)

Qy 4 TtpProTtpTtpProTtpArg 10
 Db 623 TGGCCCTGGTGGCCCTGGCGA 603

RESULT 11

US-09-252-991A-9281
 ; Sequence 9281, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 9281
 ; LENGTH: 1521
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-9281

Alignment Scores:

Pred. No.: 57.1 Length: 1521
 Score: 63.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.26% Indels: 0
 DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x US-09-252-991A-9281 (1-1521)

Qy 4 TtpProTtpTtpProTtpArg 10
 Db 992 TGGCCCTGGTGGCCCTGGCGA 1012

RESULT 12

US-08-793-035-6/c
 ; Sequence 6, Application US/08793035
 ; Patent No. 6011201
 ; GENERAL INFORMATION:
 ; APPLICANT: Slabas, Antoni R.
 ; APPLICANT: White, Andrew
 ; APPLICANT: Chase, Dianne
 ; APPLICANT: Elborough, Keiran
 ; APPLICANT: Fentem, Phillip A.
 ; TITLE OF INVENTION: B-Ketoacyl ACP Reductase Genes From
 ; TITLE OF INVENTION: Brassica Napus
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: US
 ; ZIP: 77210-4433
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/793,035
 ; FILING DATE: 28-JUL-1997
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9414622.2
 ; FILING DATE: 20-JUL-1994

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB95/01678
 ; FILING DATE: 17-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kammerer, Patricia A.
 ; REGISTRATION NUMBER: 29,775
 ; REFERENCE/DOCKET NUMBER: MOBT:132
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713.787.1400
 ; TELEFAX: 713.787.1440
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 758 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-793-035-6

Alignment Scores:
 Pred. No.: 47.3 Length: 758
 Score: 61.00 Matches: 7
 Percent Similarity: 88.89% Conservative: 1
 Best Local Similarity: 77.78% Mismatches: 1
 Query Match: 70.93% Indels: 0
 DB: 3 Gaps: 0

US-09-444-281-36 (1-12) x US-08-793-035-6 (1-758)

Qy 1 lileuArgTtpProTtpTtp 9
 Db 451 GTCCCTGGATGGAGGTGGGCCATGG 425

RESULT 13

US-09-118-554-59
 ; Sequence 59, Application US/09118554A
 ; Patent No. 6365348
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Xu, Jiangchun
 ; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
 ; TITLE OF INVENTION: METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.450C1
 ; CURRENT APPLICATION NUMBER: US/09/118,554A
 ; CURRENT FILING DATE: 1998-07-17
 ; EARLIER APPLICATION NUMBER: 08/988,255
 ; EARLIER FILING DATE: 1997-12-24
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 59
 ; LENGTH: 382
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-118-554-59

Alignment Scores:
 Pred. No.: 39.6 Length: 382
 Score: 59.00 Matches: 7
 Percent Similarity: 87.50% Conservative: 0
 Best Local Similarity: 87.50% Mismatches: 1
 Query Match: 68.60% Indels: 0
 DB: 4 Gaps: 0

US-09-444-281-36 (1-12) x US-09-118-554-59 (1-382)

Qy 2 leuArgTtpProTtpTtp 9
 |||||

Job time : 48.08 secs

```
Db      155 CTTGCTGGCCTTGGTGGAGCTGG 178
RESULT 14
US-09-118-627-59
; Sequence 59, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
; TITLE OF INVENTION: AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C1
; CURRENT APPLICATION NUMBER: US/09/118,627A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-118-627-59
Alignment Scores:
Pred. No.:      39.6      Length:      382
Score:          59.00     Matches:      7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match:    68.60% Indels: 0
DB:             4        Gaps: 0
US-09-444-281-36 (1-12) x US-09-118-627-59 (1-382)
Qy      2 LeuArgTrpProTIpTrpProTIp 9
|||||
Db      155 CTTGCTGGCCTTGGTGGAGCTGG 178
RESULT 15
US-09-602-877A-59
; Sequence 59, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-59
Alignment Scores:
Pred. No.:      39.6      Length:      382
Score:          59.00     Matches:      7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match:    68.60% Indels: 0
DB:             4        Gaps: 0
US-09-444-281-36 (1-12) x US-09-602-877A-59 (1-382)
Qy      2 LeuArgTrpProTIpTrpProTIp 9
|||||
Db      155 CTTGCTGGCCTTGGTGGAGCTGG 178
```

Search completed: May 11, 2004, 16:58:19

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:35:33 ; Search time 259.68 Seconds
(without alignments)
209.334 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWFVFWRRK 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09444281/runat_07052004_171139_2611/app_query.fasta_1.398
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl6sum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09444281 @CGN 1.1 271 @runat 07052004 171139_2611
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	Query				

1	86	100.0	72	16	US-10-395-896-32	Sequence 32, Appl
2	86	100.0	121	16	US-10-395-896-28	Sequence 28, Appl
c 3	86	100.0	121	16	US-10-395-896-62	Sequence 62, Appl
4	86	100.0	123	16	US-10-395-896-29	Sequence 29, Appl
5	86	100.0	139	16	US-10-395-896-31	Sequence 31, Appl
6	86	100.0	166	16	US-10-395-896-30	Sequence 30, Appl
7	81	94.2	114	16	US-10-395-896-64	Sequence 64, Appl
c 8	81	94.2	114	16	US-10-395-896-65	Sequence 65, Appl
9	74	86.0	68	9	US-09-309-652-6	Sequence 6, Appl
10	70	81.4	550	15	US-10-076-816-60	Sequence 60, Appl
11	70	81.4	6446	10	US-09-962-527-5	Sequence 5, Appl
12	67	77.9	207	15	US-10-252-773-25	Sequence 25, Appl
13	65	75.6	39	15	US-10-252-773-15	Sequence 15, Appl
c 14	65	75.6	39	15	US-10-252-773-19	Sequence 19, Appl
15	65	75.6	54	15	US-10-252-773-17	Sequence 17, Appl
16	64	74.4	557	13	US-10-027-632-53929	Sequence 53929, A
17	64	74.4	557	13	US-10-027-632-321717	Sequence 321717, A
18	64	74.4	557	16	US-10-027-632-53929	Sequence 53929, A
19	64	74.4	557	16	US-10-027-632-321717	Sequence 321717, A
20	63.5	73.8	571	13	US-10-425-114-14181	Sequence 14181, A
21	63	73.3	1437	9	US-09-815-242-7766	Sequence 7766, Ap
c 22	63	73.3	1437	9	US-09-815-242-7716	Sequence 9716, Ap
23	63	73.3	1908	13	US-10-282-122A-39517	Sequence 39517, A
c 24	63	73.3	2403	15	US-10-156-761-4600	Sequence 4600, Ap
25	63	73.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
c 26	61.5	71.5	1224	16	US-10-363-493-40678	Sequence 40678, A
c 27	61	70.9	496	13	US-10-424-599-92875	Sequence 92875, A
c 28	61	70.9	577	15	US-10-029-386-12845	Sequence 12845, A
c 29	61	70.9	1526	13	US-10-424-599-17644	Sequence 17644, A
c 30	61	70.9	1825	16	US-10-116-375-321	Sequence 321, App
c 31	61	70.9	2218	10	US-09-820-790-1	Sequence 1, Appli
32	61	70.9	34570	12	US-09-997-722-157	Sequence 157, App
c 33	60.5	70.3	412	9	US-09-983-965-2138	Sequence 2138, A
c 34	60.5	70.3	1038	13	US-10-425-114-23253	Sequence 23253, A
c 35	60.5	70.3	1072	13	US-10-425-114-30359	Sequence 30359, A
c 36	60	69.8	429	13	US-10-424-599-1131	Sequence 1131, Ap
37	60	69.8	837	13	US-10-424-599-17642	Sequence 17642, A
38	60	69.8	1440	13	US-10-282-122A-27681	Sequence 27681, A
c 39	60	69.8	1596	13	US-10-425-114-14091	Sequence 14091, A
40	59.5	69.2	733	13	US-10-425-114-4367	Sequence 4367, Ap
41	59	68.6	50	16	US-10-131-827-2961	Sequence 2961, Ap
42	59	68.6	335	13	US-10-085-783A-13208	Sequence 13208, A
c 43	59	68.6	335	16	US-10-242-535A-13208	Sequence 13208, A
44	59	68.6	382	9	US-09-745-288-59	Sequence 59, Appl
45	59	68.6	382	13	US-10-453-919-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-10-395-896-32
; Sequence 32, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; TITLE OF INVENTION: ANTI-INFECTION PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;
; FEATURE:
; OTHER INFORMATION: Exemplary "template"
US-10-395-896-32

Alignment Scores:
Pred. No.: 72 Length: 72
Score: 0.00445 Matches: 12
Percent Similarity: 86.00 Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-36 (1-12) x US-10-395-896-32 (1-72)

QY 1 IleLeuArgTrpProTrrpProTrrpArgArgLys 12
Db 5 ATTCTGCGTTGGCGGTGGTGGCGGTGGCGGTGGCGAAA 40

RESULT 2
US-10-395-896-28
; Sequence 28, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; TITLE OF INVENTION: ANTI-INFECTIVE PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Exemplary "template"
US-10-395-896-28

Alignment Scores:
Pred. No.: 121 Length: 121
Score: 0.00712 Matches: 12
Percent Similarity: 86.00 Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-36 (1-12) x US-10-395-896-28 (1-121)

QY 1 IleLeuArgTrpProTrrpProTrrpArgArgLys 12
Db 48 ATTCTGCGTTGGCGGTGGTGGCGGTGGCGGTGGCGAAA 83

RESULT 3
US-10-395-896-62/c
; Sequence 62, Application US/10395896
; Publication No. US20030219854A1
; GENERAL INFORMATION:
; APPLICANT: Guarna, Maria Marta
; APPLICANT: Chen, Yuchen
; APPLICANT: Cory, Robert
; APPLICANT: Brinkman, Jacqui
; APPLICANT: Cabralda, Jennifer
; APPLICANT: Metlitskaia, Luba
; APPLICANT: Suleman, Dinar
; TITLE OF INVENTION: METHODS FOR PRODUCING MODIFIED
; TITLE OF INVENTION: ANTI-INFECTIVE PEPTIDES
; FILE REFERENCE: 660081.421
; CURRENT APPLICATION NUMBER: US/10/395,896
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Exemplary "template"
US-10-395-896-29

Alignment Scores:
Pred. No.: 123 Length: 123
Score: 0.00723 Matches: 12
Percent Similarity: 86.00 Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-444-281-36 (1-12) x US-10-395-896-29 (1-123)

QY 1 IleLeuArgTrpProTrrpProTrrpArgArgLys 12
Db 26 ATTCTGCGTTGGCGGTGGTGGCGGTGGCGGTGGCGAAA 61

RESULT 5
US-10-395-896-31
; Sequence 31, Application US/10395896
; Publication No. US20030219854A1
```

US-10-395-896-31

US-10-395-896-65

Alignment Scores: 114
Pred. No.: 0.0278
Score: 81.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 94.19%
Indels: 0
Gaps: 0

US-09-444-281-36 (1-12) x US-10-395-896-65 (1-114)

QY 1 lleuArgTrpProTrpArgArg 11
DB 85 ATTCTGGTGGCGGTGGCGGTGGCGTGC 53

RESULT 9

US-09-909-652-6
; Sequence 6, Application US/09909652
; Patent No. US2002002537A1
; GENERAL INFORMATION:
; APPLICANT: Kairos Scientific, Inc.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; APPLICANT: Youvan, Douglas C.
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR GENERATING
; TITLE OF INVENTION: AND SCREENING COMPOUNDS THAT AFFECT CELL VIABILITY
; FILE REFERENCE: 22346-7001
; CURRENT APPLICATION NUMBER: US/09/909,652
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/219,179
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 68
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid encoding ubiquitin indolicidin fusion
; OTHER INFORMATION: protein fragment
US-09-909-652-6

Alignment Scores: 68
Pred. No.: 0.127
Score: 74.00
Percent Similarity: 100.00%
Best Local Similarity: 90.00%
Query Match: 86.05%
Indels: 0
Gaps: 0

US-09-444-281-36 (1-12) x US-09-909-652-6 (1-68)

QY 2 leuArgTrpProTrpArgArg 11
DB 28 TTGAATGGCGTGGCGGTGGCGTGC 57

RESULT 10

US-10-076-816-60
; Sequence 60, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Bos taurus
US-10-076-816-60

Alignment Scores: 550
Pred. No.: 2.61
Score: 70.00
Percent Similarity: 100.00%
Best Local Similarity: 88.89%
Query Match: 81.40%
Indels: 0
Gaps: 0

US-09-444-281-36 (1-12) x US-10-076-816-60 (1-550)

QY 3 ArgTrpProTrpProTrpArg 11
DB 415 AAATGCCCATGGTGGCGGTGGCGGAGA 441

RESULT 11

US-09-962-527-5
; Sequence 5, Application US/09962527
; Publication No. US20030049813A1
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; HOLTZ, R. BARRY
; McCULLOCH, MICHAEL
; TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,527
; FILING DATE: 24-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: 10-march-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 5;
US-09-962-527-5

Alignment Scores:
Pred. No.: 24.2 Length: 6446
Score: 70.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 81.40% Indels: 0
DB: 10 Gaps: 0

US-09-444-281-36 (1-12) x US-09-962-527-5 (1-6446)
QY 3 ArgTrpProTTrpProTTrpArg 11
Db 6213 AAGUGGCCUUGGGCCCAUGCGCCGA 6239

RESULT 12
US-10-252-773-25
; Sequence 25, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA PCRIL
US-10-252-773-25

Alignment Scores:
Pred. No.: 2.53 Length: 207
Score: 67.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 77.91% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-36 (1-12) x US-10-252-773-25 (1-207)
QY 1 IleuArgTrpProTTrpProTTrpArg 10
Db 163 ATTAGGAGTGCCTTGGTGGCCTTGGA 192

RESULT 13
US-10-252-773-15
; Sequence 15, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-252-773-15

Alignment Scores:
Pred. No.: 1.3 Length: 53
Score: 65.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 75.58% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-36 (1-12) x US-10-252-773-15 (1-53)
QY 3 ArgTrpProTTrpProTTrpArg 10
Db 50 AGATGGCCTTGGTGGCCTTGGA 27

RESULT 14
US-10-252-773-17/c
; Sequence 17, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 53
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-252-773-17

Alignment Scores:
Pred. No.: 1.3 Length: 53
Score: 65.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 75.58% Indels: 0
DB: 15 Gaps: 0

US-09-444-281-36 (1-12) x US-10-252-773-17 (1-53)
QY 3 ArgTrpProTTrpProTTrpArg 10
Db 50 AGATGGCCTTGGTGGCCTTGGA 27

RESULT 15
```

US-10-252-773-19/c
; Sequence 19, Application US/10252773
; Publication No. US2003013183A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, QUNIGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252,773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106,373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106,573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 54
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Oligonucleotide
US-10-252-773-19
Alignment Scores:
Pred. No.: 1.32 Length: 54
Score: 65.00 Matches: 7
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 0
Query Match: 75.58% Indels: 0
DB: 15 Gaps: 0
US-09-444-281-36 (1-12) x US-10-252-773-19 (1-54)
Qy 3 ArgTIpProTIpTIpProTIpArg 10
Db 51 AGATGGCCTTGGTGGCCTTGGAAA 28

Search completed: May 11, 2004, 17:07:35
Job time : 262.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 11, 2004, 14:00:48 ; Search time 1588.8 Seconds
(without alignments)
225.545 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPPWPRK 12

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -US09444281/runat 07052004 171138 2579/app query.fasta_1.398
-Q/cgn2 1/USPTO.spool -SURFEX=0.1 -LOOPCL=0 -LOOPEXT=0
-DB=EST -QFMT=fastap -SURFEX=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -NATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09444281 @CN 1 1 3596 @runat 07052004 171138 2579 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estma:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	87.2	522	14	CF479395
C 2	69	80.2	446	9	AU198144
C 3	69	80.2	448	9	AU198162
C 4	69	80.2	539	14	CF326637
C 5	69	80.2	550	14	CF326444
C 6	69	80.2	559	14	CF325873
C 7	69	80.2	578	9	AU082117
C 8	69	80.2	595	14	CF326279
C 9	69	80.2	661	14	CB681162
10	69	80.2	742	12	BI698060
11	69	80.2	946	13	BQ931041
12	69	80.2	959	13	BQ877964
13	69	80.2	984	13	BQ715181
14	69	80.2	1011	13	BQ715516
15	69	80.2	2675	11	BC051074
16	69	80.2	3114	11	BC034057
17	69	80.2	3751	11	AK080277
18	69	80.2	3782	11	AK041860
19	68	79.1	249	14	CF850934
20	68	79.1	349	12	BJ476852
21	68	79.1	369	12	BJ474341
C 22	68	79.1	381	12	BJ471810
C 23	68	79.1	411	12	BJ473016
C 24	68	79.1	420	9	AV933841
C 25	68	79.1	440	12	BJ471261
C 26	68	79.1	446	9	AU089922
C 27	68	79.1	450	14	CF862143
28	68	79.1	464	9	AV935002
29	68	79.1	471	12	BJ475478
30	68	79.1	472	12	BJ475115
31	68	79.1	472	12	BJ476710
32	68	79.1	473	9	AV937233
33	68	79.1	474	12	BJ473880
C 34	68	79.1	475	9	AU089934
35	68	79.1	489	12	BJ477184
36	68	79.1	500	12	BJ475153
C 37	68	79.1	501	12	BJ470669
38	68	79.1	504	12	BJ475306
C 39	68	79.1	506	12	BJ471761
40	68	79.1	506	12	BJ473882
C 41	68	79.1	509	12	BJ471258
C 42	68	79.1	515	12	BJ469760
C 43	68	79.1	530	12	BJ472623
44	68	79.1	531	12	BJ475836
C 45	68	79.1	532	9	AV932159

ALIGNMENTS

RESULT 1
CF479395
LOCUS
DEFINITION
RTMW3_23_A01.g1_A022 Well-watered loblolly pine roots WM3 Pinus
taeda cDNA clone RTMW3_23_A01_A022 5', mRNA sequence.
ACCESSION
CF479395
VERSION
CF479395.1 GI:34508264
KEYWORDS
EST.
SOURCE
Pinus taeda (loblolly pine)
ORGANISM
Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE
1 (bases 1 to 522)

AUTHORS

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., and Neale, D.
 An EST database from well-watered loblolly pine (*Pinus taeda*) roots
 Unpublished (2003)
 Other ESTs: RTW3 23 A01.b1 A022
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
 Seq primer: JENREV (CAGGACACGCTATGACC).

FEATURES

source
 1..522
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /strain="CCIONES"
 /db_xref="taxon:3352"
 /clone="RTW3 23 A01 A022"
 /lab_host="DH10B-TI phage-resistant E. coli"
 /clone_lib="Well-watered loblolly pine roots W3"
 /notes="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (*Pinus taeda*) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Alignment Scores:
 Pred. No.: 203 Length: 522
 Score: 75.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 90.00% Mismatches: 0
 Query Match: 87.21% Indels: 0
 DB: 14 Gaps: 0

US-09-444-281-36 (1-12) x CF479395 (1-522)

QY 3 ArgTTPProTTPProTTPArgArglys 12

Db 287 AGGTGGCCCTGGTGGCCCTGGCGGCGAGG 316

RESULT 2

AU198144/c
 LOCUS AU198144 Rice green shoot Oryza sativa (japonica cultivar-group)
 DEFINITION cDNA clone S15951, mRNA sequence.

ACCESSION AU198144

VERSION AU198144.1 GI:14714211

KEYWORDS

SOURCE EST.

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 446)
 Sasaki, T. and Yamamoto, K.
 Rice cDNA from green shoot (2001)
 Unpublished (2001)

JOURNAL

CONTACT: Takuji Sasaki

National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
 PROJECT = 'RGP'.

FEATURES

source
 1..446
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S15951"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

ORIGIN

Alignment Scores:
 Pred. No.: 687 Length: 446
 Score: 90.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 80.23% Indels: 0
 DB: 9 Gaps: 0

US-09-444-281-36 (1-12) x AU198144 (1-446)

QY 3 ArgTTPProTTPProTTPArgArglys 12

Db 348 CGCTGGCCCTGGTGGCCCTGGACGGCGCG 319

RESULT 3

AU198162/c
 LOCUS AU198162 Rice green shoot Oryza sativa (japonica cultivar-group)

DEFINITION cDNA clone S16019, mRNA sequence.

ACCESSION AU198162

VERSION AU198162.1 GI:14714231

KEYWORDS

SOURCE EST.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 448)
 Sasaki, T. and Yamamoto, K.
 Rice cDNA from green shoot (2001)
 Unpublished (2001)

AUTHORS

CONTACT: Takuji Sasaki

JOURNAL

National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki

COMMENT

305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
 PROJECT = 'RGP'.

FEATURES

source
 1..448
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S16019"
 /clone_lib="Rice green shoot"
 /note="Green shoot (8 days old)"

ORIGIN

Alignment Scores:
 Pred. No.: 691 Length: 448
 Score: 69.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 80.23% Indels: 0
 DB: 9 Gaps: 0

US-09-444-281-36 (1-12) x AUL98162 (1-448)

QY 3 ArgTtpProTtpTtpProTtpArgArgLys 12
Db 311 CGCTGGCCTTGTGGCCCTGGACGGCGG 282

RESULT 4

CF326637/c
LOCUS CF326637 539 bp mRNA linear EST 18-AUG-2003
DEFINITION JMT1--06-J20-g1 AtJMT-overexpressing transgenic rice lambda phage cDNA library (JMT1) Oryza sativa cDNA clone JMT1--06-J20, mRNA sequence.

ACCESSION CF326637
VERSION CF326637.1 GI:33801530
KEYWORDS EST.
SOURCE Oryza sativa

ORGANISM Oryza sativa

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 539)

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.myongji.ac.kr.

FEATURES

source

1..539
Location/Qualifiers
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="JMT1--06-J20"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="AtJMT-overexpressing transgenic rice lambda phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was prepared from Arabidopsis Jasmonate Carboxyl methyltransferase overexpression line."

ORIGIN

Alignment Scores:
Pred. No.: 860 Length: 539
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 14 Gaps: 0

US-09-444-281-36 (1-12) x CF326637 (1-539)

QY 3 ArgTtpProTtpTtpProTtpArgArgLys 12
Db 326 CGCTGGCCTTGTGGCCCTGGACGGCGG 297

RESULT 5

CF326444/c
LOCUS CF326444 550 bp mRNA linear EST 18-AUG-2003
DEFINITION JMT1--06-B08-g1 AtJMT-overexpressing transgenic rice lambda phage cDNA library (JMT1) Oryza sativa cDNA clone JMT1--06-B08, mRNA sequence.

ACCESSION CF326444

VERSION CF326444.1 GI:33801141

KEYWORDS EST.

SOURCE Oryza sativa

ORGANISM Oryza sativa

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 550)

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.myongji.ac.kr.

FEATURES

source

1..550
Location/Qualifiers
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="JMT1--06-B08"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="AtJMT-overexpressing transgenic rice lambda phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was prepared from Arabidopsis Jasmonate Carboxyl methyltransferase overexpression line."

ORIGIN

Alignment Scores:
Pred. No.: 881 Length: 550
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 14 Gaps: 0

US-09-444-281-36 (1-12) x CF326444 (1-550)

QY 3 ArgTtpProTtpTtpProTtpArgArgLys 12

Db 352 CGCTGGCCTTGTGGCCCTGGACGGCGG 323

RESULT 6

CF325873/c

LOCUS CF325873

DEFINITION JMT1--04-H09-g1 AtJMT-overexpressing transgenic rice lambda phage cDNA library (JMT1) Oryza sativa cDNA clone JMT1--04-H09, mRNA sequence.

ACCESSION CF325873

VERSION CF325873.1 GI:33800014

KEYWORDS EST.

SOURCE Oryza sativa

ORGANISM Oryza sativa

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 559)

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnam@gbio.com, bhnam@bio.myongji.ac.kr.

FEATURES

source
1. .559
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="JMT1--04-H09"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="AtUMT-overexpressing transgenic rice lambda
phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5',
end with EcoRI and 3' end with XhoI site. mRNA was
prepared from Arabidopsis Jasmonate Carboxyl
methyltransferase overexpression line."

-ORIGIN

Alignment Scores:
Pred. No.: 898 Length: 559
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 14 Gaps: 0

US-09-444-281-36 (1-12) x CF325873 (1-559)

QY 3 ArgTrpProTrpTrpProTrpArgArgLys 12

DB 324 CGCTGGCCTTGGTGGCCCTGGACGCGGG 295

RESULT 7

LOCUS AU082117/c 578 bp mRNA linear EST 02-APR-2002
DEFINITION AU082117 Rice panicle at ripening stage Oryza sativa (japonica
cultivar-group) cDNA clone E11611, mRNA sequence.

ACCESSION AU082117

VERSION AU082117.1 GI:6727452

KEYWORDS EST

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 578)

AUTHORS Sasaki, T. and Yamamoto, K.

TITLE Rice cDNA from panicle at ripening stage (2000)

JOURNAL Unpublished (2000)

COMMENT Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program, Kamondai 2-1-2, Tsukuba, Ibaraki

305-8602, Japan

Tel: 81-298-38-7441

Fax: 81-298-38-7468

Email: tsasaki@affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/

PROJECT = RGP.

FEATURES

source
1. .578
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E11611"
/dev_stage="ripening stage"
/clone_lib="Rice panicle at ripening stage"
/note="Organ: panicle; Rice cDNA from panicle at ripening

ORIGIN

Alignment Scores:
Pred. No.: 935 Length: 578
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 9 Gaps: 0

US-09-444-281-36 (1-12) x AU082117 (1-578)

QY 3 ArgTrpProTrpTrpProTrpArgArgLys 12

DB 345 CGCTGGCCTTGGTGGCCCTGGACGCGGG 316

RESULT 8

LOCUS CF326279/c

DEFINITION JMT1--05-K03 g1 AtUMT-overexpressing transgenic rice lambda phage
cDNA library (JMT1) Oryza sativa cDNA clone JMT1--05-K03, mRNA
sequence.

ACCESSION CF326279

VERSION CF326279.1 GI:33800815

KEYWORDS EST

SOURCE Oryza sativa

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 595)

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,

Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

CONTACT: Nahm B.H.

of Bioscience and Bioinformatics, GreenGene Biotech Inc.; Division

Genomics and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnam@gbio.com, bhnam@bio.myongji.ac.kr.

FEATURES

source
1. .595
Location/Qualifiers
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="JMT1--05-K03"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E.coli SOLR"
/clone_lib="AtUMT-overexpressing transgenic rice lambda
phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5',
end with EcoRI and 3' end with XhoI site. mRNA was
prepared from Arabidopsis Jasmonate Carboxyl
methyltransferase overexpression line."

ORIGIN

Alignment Scores:
Pred. No.: 967 Length: 595
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 14 Gaps: 0

US-09-444-281-36 (1-12) x CF326279 (1-595)

QY 3 ArgTrpProTrpTrpProTrpArgArgLys 12

```

Db      322  CGCTGGCTTGGTGGCCCTGACGCGCGG 293
|||||
RESULT 9
CB681162/c
LOCUS   661 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf07B03.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEf07B03 5', mRNA sequence.
ACCESSION CB681162
VERSION   1 GI:29684887
KEYWORDS EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 661)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: B column: 03
Seq primer: gta aaa cga cgg cca gtg.
FEATURES
source
location/Qualifiers
1..661
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEf07B03"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEf"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"
ORIGIN
Alignment Scores:
Pred. No.: 1..61 Length: 661
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 14 Gaps: 0
US-09-444-281-36 (1-12) x CB681162 (1-661)
Qy 3 ArgTTPProTTPProTTPArgArgLys 12
|||||
Db 341 CGCTGGCTTGGTGGCCCTGACGCGCGG 312
|||||
RESULT 10
BI698060
LOCUS 742 bp mRNA linear EST 18-SEP-2001
DEFINITION 603344220F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5371828 5',
mRNA sequence.
ACCESSION BI698060
VERSION BI698060.1 GI:15660699
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 742)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11946 row: h column: 05
High quality sequence stop: 742.
location/Qualifiers
1..742
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5371828"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
ORIGIN
Alignment Scores:
Pred. No.: 1..26e+03 Length: 742
Score: 69.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.23% Indels: 0
DB: 12 Gaps: 0
US-09-444-281-36 (1-12) x BI698060 (1-742)
Qy 3 ArgTTPProTTPProTTPArgArgLys 12
|||||
Db 222 AGGTGGCTTGGTGGCCCTGAGGTCTCGG 251
|||||
RESULT 11
BI9311041
LOCUS 946 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT 8933817 NCI_CGAP_Mam2 Mus musculus cDNA clone
IMAGE:6489249 5', mRNA sequence.
ACCESSION BI9311041
VERSION BI9311041.1 GI:22346072
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

http://image.llnl.gov
 Plate: LLAM14037 row: c column: 10
 High quality sequence stop: 539.

FEATURES
 source
 1. .946
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:6489249"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Man2"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:
 Pred. No.: 1.68e+03 Length: 946
 Score: 69.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 80.23% Indels: 0
 DB: 13 Gaps: 0

US-09-444-281-36 (1-12) x BQ931041 (1-946)

QY 3 ArgTTPProTTPProTTPArgArgLys 12
 |||||
 DB 76 AGTGGCCCTTGGTGGCCCTGGAGGCTCGG 105

RESULT 12

BQ877964
 LOCUS BQ877964 959 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOURT_8712713 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6307097
 5', mRNA sequence.
 ACCESSION BQ877964
 VERSION BQ877964.1 GI:22269972
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 959)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13724 row: a column: 18
 High quality sequence stop: 598.

FEATURES
 source

1. .959
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6307097"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_129"
 /note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dT. Average insert size 2.2 kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH_MGC

ORIGIN

Alignment Scores:
 Pred. No.: 1.7e+03 Length: 959
 Score: 69.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 80.23% Indels: 0
 DB: 13 Gaps: 0

US-09-444-281-36 (1-12) x BQ877964 (1-959)

QY 3 ArgTTPProTTPProTTPArgArgLys 12
 |||||
 DB 199 AGTGGCCCTTGGTGGCCCTGGAGGCTCGG 228

RESULT 13

BQ715181
 LOCUS BQ715181 984 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8474869 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310462
 5', mRNA sequence.
 ACCESSION BQ715181
 VERSION BQ715181.1 GI:21854080
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 984)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13732 row: m column: 23
 High quality sequence stop: 503.

FEATURES
 source

1. .984
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6310462"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_129"
 /note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dT. Average insert size 2.2 kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
 Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.76e+03 Length: 984
 Score: 69.00 Matches: 8
 Percent Similarity: 90.00% Conservative: 1
 Best Local Similarity: 80.00% Mismatches: 1
 Query Match: 80.23% Indels: 0
 DB: 13 Gaps: 0

US-09-444-281-36 (1-12) x BQ715181 (1-984)

QY 3 ArgTTPProTTPProTTPArgArgLys 12
 |||||
 DB 415 AGTGGCCCTTGGTGGCCCTGGAGGCTCGG 444

RESULT 14

BQ715516
 LOCUS BQ715516 1011 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT 8294996 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6309470
 5', mRNA sequence.

ACCESSION BQ715516
 VERSION BQ715516.1 GI:21854415

KEYWORDS EST.

SOURCE BQ715516.1 GI:21854415

ORGANISM Mus musculus (house mouse)

REFERENCE

1 (bases 1 to 1011)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

JOURNAL NIH-MGC http://mgc.nci.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

CONTACT Robert Strausberg, Ph.D.

EMAIL: cgaabs-r@mail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.

cDNA Library Preparation: ResGen, Invitrogen Corp.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13730 row: d column: 15

High quality sequence stop: 554.

Location/Qualifiers

1..1011

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6309470"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_129"

/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dt. Average insert size 2.2 kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
 Library."

ORIGIN

Alignment Scores:

Pred. No.: 1

Score: 1.81e+03 Length: 1011

Percent Similarity: 69.00% Matches: 8

Best Local Similarity: 90.00% Conservative: 1

Query Match: 80.00% Mismatches: 1

DB: 13 Indels: 0 Gaps: 0

US-09-444-281-36 (1-12) x BQ715516 (1-1011)

QY 3 ArgTTPProTTPProTTPArgArgLys 12

DB 198 AGTGGCCCTTGGTGGCCCTGGAGTCTCG 227

BC051074

Mus musculus elastin, mRNA (cDNA clone IMAGE:6390348), with

apparent retained intron.

ACCESSION BC051074

VERSION BC051074.1 GI:30048428

KEYWORDS HTC.

SOURCE BQ715516.1 GI:21854415

ORGANISM Mus musculus (house mouse)

REFERENCE

1 (bases 1 to 2675)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,

Schretz J.E., Brownstein M.J., Uedin T.B., Toshiyuki S.,

Carinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,

Abrams R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,

McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,

Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Faney J., Helton E., Kettman M., Madan A., Rodriguez S.,

Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,

Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,

Bickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

Butterfield Y.S., Krzywinski M.I., Skalska U., Smalios D.E.,

Scherch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2675)

Strausberg R.

Direct Submission

Submitted (11-APR-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.

cDNA Library Preparation: ResGen, Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Contact: nisc.mgc@nih.gov

Web site: http://www.nisc.nih.gov/

Akhter N., Ayèle K., Beckstrom-Sternberg S.M., Benjamin B.,

Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,

Dietrich N.L., Granito S., Guan X., Gupta J., Haghighi P.,

Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,

Maduro Q.L., Masiello C., Maskeri B., Mastrian S.D., McCloskey J.C.,

McDowell J., Pearson R., Stantripop S., Thomas P.J., Touchman J.W.,

Tsurgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,

Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAK Plate: 108 Row: 1 Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6679636

This clone has the following problem: retained intron.

Location/Qualifiers

1..2675

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6390348"

/tissue_type="olfactory epithelium, neonatal mouse,

C57Bl/6"

/clone_lib="NIH_MGC_129"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 5.74e+03 Length: 2675

Score: 69.00 Matches: 8

Percent Similarity: 90.00% Conservative: 1

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 80.23% Indels: 0

DB: 11 Gaps: 0

US-09-444-281-36 (1-12) x BC051074 (1-2675)

Qy 3 ArgTIpProTIpTIpProTIpArgArgLys 12
Db 1112 AGGTGGCCTTGGTGGCCCTGGAGGTCGCG 1141

Search completed: May 11, 2004, 16:56:35
Job time : 1591.8 secs